PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) **International Patent Classification** 6 :
C12N 15/15, C07K 14/81, C12N 1/20, 15/62, C12Q 1/37 — A1

(11) **International Publication Number:** WO 95/02055

(43) **International Publication Date:** 19 January 1995 (19.01.95)

(21) **International Application Number:** PCT/US94/07779

(22) **International Filing Date:** 8 July 1994 (08.07.94)

(30) **Priority Data:**
08/089,248 — 9 July 1993 (09.07.93) — US

(71)(72) **Applicants and Inventors:** LASKOWSKI, Michael [US/US]; 222 E. Navajo, W. Lafayette, IN 47906 (US). ANDERSON, Stephen [US/US]; 158 Springdale Street, Princeton, NJ 08540 (US).

(74) **Agents:** BEHR, Omri, M. et al.; Behr and Adams, 325 Pierson Avenue, Edison, NJ 08837-3123 (US).

(81) **Designated States:** AU, BB, BG, BR, BY, CA, CN, CZ, FI, HU, JP, KP, KR, KZ, LK, LV, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SK, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) **Title:** NOVEL PROTEIN INHIBITORS OF SERINE PROTEINASES (E.G. FURIN) DERIVED FROM TURKEY OVOMUCOID THIRD DOMAIN

(57) **Abstract**

Novel protein inhibitors of serine proteinases incorporating immediately adjacent to the inhibitor reactive site peptide bond a consensus sequence from substrates of a particular proteinase. The inhibitors include Kazal family inhibitors of subtilisin-like proteinases, and more particularly analogs of turkey ovomucoid third domain protein (6-56) such as A15R-T17K-L18R. The inhibitors inhibit the proteinase activity of enzymes such as human furin. There are further provided polynucleotides comprising one or more sequences of nucleotide bases collectively encoding the amino acid sequence of one of said turkey ovomucoid third domain peptide inhibitors; an expression vector comprising such a polynucleotide; an organism transformed with such an expression vector; and methods of synthesizing such an inhibitor and of inhibiting a serine proteinase.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

INTERNATIONAL SEARCH REPORT

Int Jonal Application No
PCT/US 94/07779

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/15 C07K14/81 C12N1/20 C12N15/62 C12Q1/37

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JOURNAL OF BIOLOGICAL CHEMISTRY., vol.266, no.17, 15 June 1991, BALTIMORE US pages 10727 - 10730 T. KOMIYAMA ET AL 'Replacement of P1 Leu18 by Glu18 in the reactive site of turkey ovomucoid third domain converts it into a strong inhibitor of Glu-specific Streptomyces griseus proteinase (GluSGP)' cited in the application	1
Y	see abstract see page 10727, left column ---	2-27
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 October 1994	08 -11- 1994
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Authorized officer Van der Schaal, C

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JOURNAL OF BIOLOGICAL CHEMISTRY., vol.267, no.23, 15 August 1992, BALTIMORE US pages 16396 - 16402 S. MOLLOY ET AL 'Human furin is a calcium-dependent serine endoprotease that recognizes the sequence Arg-X-X-Arg and efficiently cleaves anthrax toxin protective antigen' cited in the application see abstract ---	2-27
Y	EMBO JOURNAL, vol.11, no.7, 1992 pages 2407 - 2414 A. STIENEKE-GRÖBER ET AL 'Influenza virus hemagglutinin with multibasic cleavage site is activated by furin, a subtilisin-like endoprotease' cited in the application see abstract ---	2-27
X	CHEMICAL ABSTRACTS, vol. 107, no. 1, 6 July 1987, Columbus, Ohio, US; abstract no. 3101, S. JALLET ET AL 'Antiprotease targeting: altered specificity of alpha1-antitrypsin by amino acid replacement at the reactive center' page 295 ;	1
Y	see abstract & REV. FR. TRANSFUS. IMMUNO-HEMATOL., vol.29, no.4, 1986 pages 287 - 298 ---	2-27
X	JOURNAL OF BIOCHEMISTRY., vol.109, no.3, March 1991, TOKYO JP pages 377 - 382 S. KOJIMA ET AL 'Inhibition of Subtilisin BPN' by reaction site P1 mutants of Streptomyces subtilisin inhibitor'	1,2
Y	see abstract ---	3-27
X	JOURNAL OF BIOLOGICAL CHEMISTRY., vol.268, no.12, 25 April 1993, BALTIMORE US pages 8458 - 8465 L. WASLEY 'PACE/Furin can process the vitamin K-dependent pro-factor IX precursor within the secretory pathway'	1-3
1 Y	see page 8461, right column, paragraph 3 ---	4-27
2 X	EP,A,0 253 690 (TRANSGENE S.A.) 20 January 1988	1
Y	see the whole document ---	2-27

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	JOURNAL OF BIOLOGICAL CHEMISTRY., vol.268, no.20, 15 July 1993, BALTIMORE US pages 14583 - 14585 W. LU ET AL 'Arg15-Lys17-Arg18 turkey ovomucoid third domain inhibits human furin' ---	1-27
P,X	JOURNAL OF BIOLOGICAL CHEMISTRY., vol.268, no.33, 25 November 1993, BALTIMORE US pages 24887 - 24891 E. ANDERSON ET AL 'Inhibition of HIV-1 gp160-dependent membrane fusion by a furin -directed alpha1-antitrypsin variant'	1
Y	see the whole document -----	2-27

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int ional Application No
PCT/US 94/07779

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-0253690	20-01-88	FR-A-	2599752	11-12-87
		AU-B-	610500	23-05-91
		AU-A-	7362587	17-12-87
		CA-A-	1328706	19-04-94
		DE-A-	3786961	16-09-93
		DE-T-	3786961	20-01-94
		JP-A-	63079899	09-04-88
		US-A-	4973668	27-11-90
		ZA-A-	8704123	09-12-87

FIG. 1

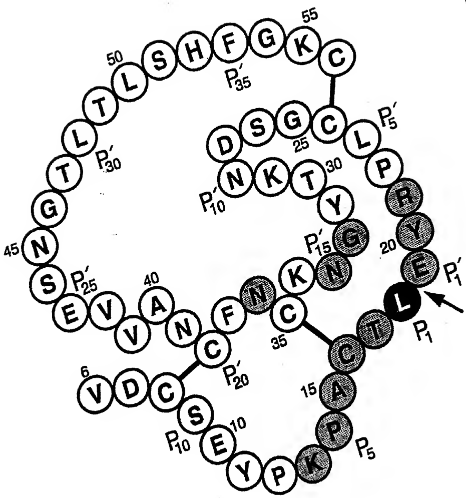
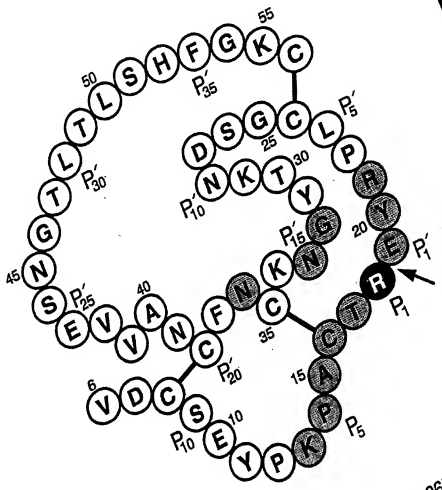
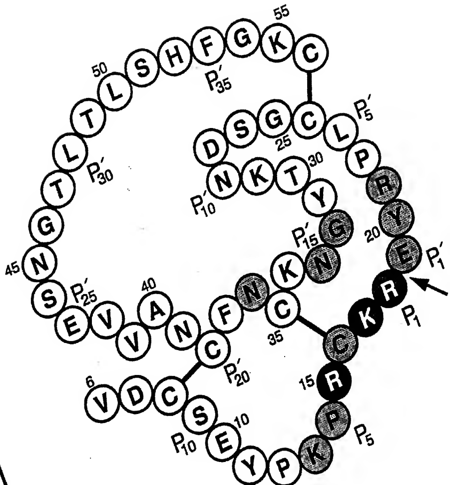

6
10
15
20
25
30
35
40
45
50
55
P1
P5
P10
P'1
P'5
P'10
P'15
P'20
P'25
P'30
P'35

FIG. 2



SUBSTITUTE SHEET (RULE 26)

FIG. 3



6
V D C S E Y P K P R C K R E Y R P L C
10
15
20
25
30
35
40
45
50
55
P_1
P_5
P_{10}
P'_1
P'_5
P'_{10}
P'_{15}
P'_{20}
P'_{25}
P'_{30}
P'_{35}

FIG. 4

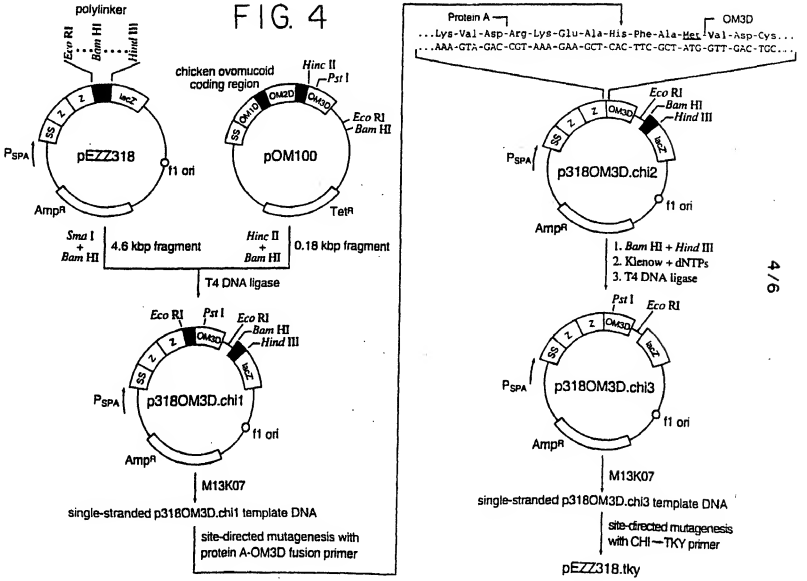


FIG. 5

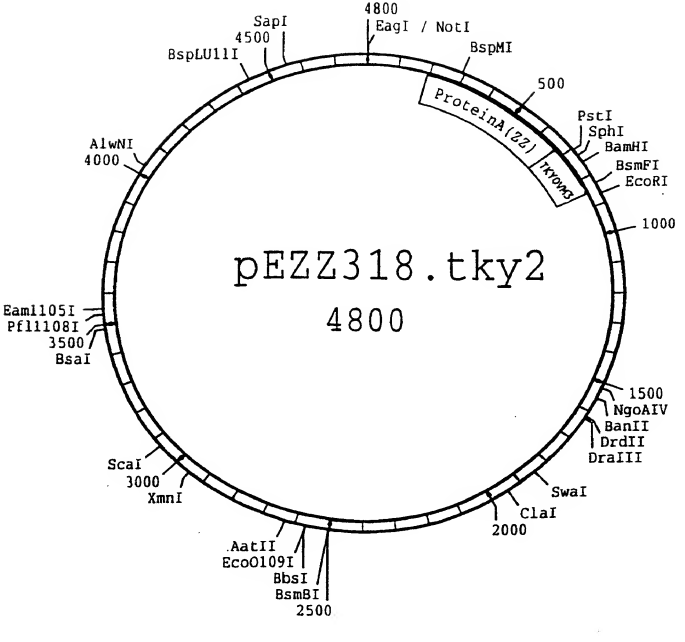
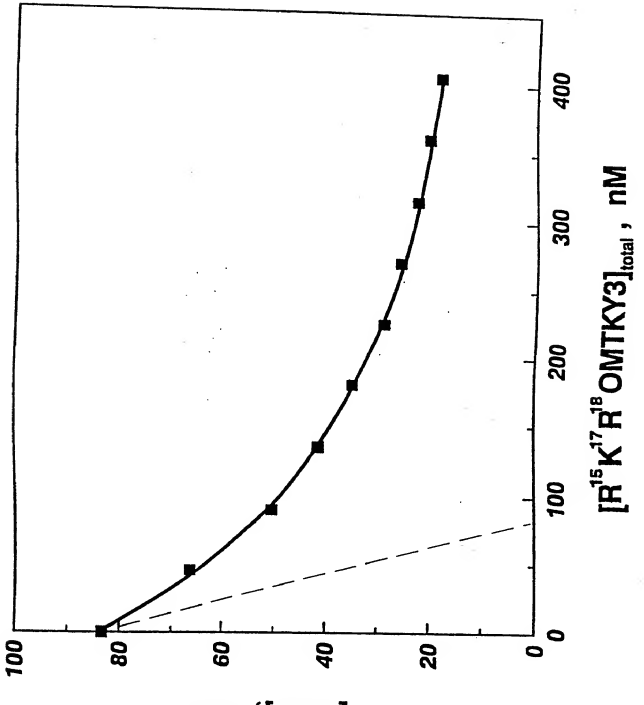

pEZZ318.tky2
4800
4800
EagI / NotI
BspMI
500
ProteinA(ZZ)
TKYORG
PstI
SphI
BamHI
BsmFI
EcoRI
1000
1500
NgoAIV
BanII
DrdII
DraIII
SwaI
ClaI
2000
AatII
Eco0109I
BbsI
BsmBI
2500
ScaI
3000
XmnI
Eam1105I
Pfl1108I
3500
BsaI
AlwNI
4000
BspLU11I
4500
SapI

FIG.6


[furin], nM
0
20
40
60
80
100
0
100
200
300
400
[R15K17R18OMTKY3]total , nM

NOVEL PROTEIN INHIBITORS OF SERINE PROTEINASES (E.G. FURIN) DERIVED FROM TURKEY OVOMUCOID THIRD DOMAIN

FIELD OF THE INVENTION

This invention concerns novel protein inhibitors of serine proteinases which incorporate immediately adjacent to the inhibitor reactive site peptide bond a con-sensus sequence from substrates of the target proteinase. The inhibitors include Kazal family inhibitors of subtilisin-like proteinases, and more particularly analogs of turkey ovomucoid third domain protein (6-56) such as A15R-T17K-L18R. The inhibitors inhibit the proteinase activity of enzymes such as human furin. There are further provided polynucleotides comprising one or more sequences of nucleotide bases collectively encoding the amino acid sequence of one of said turkey ovomucoid third domain peptide inhibitors; an expression vector comprising such a polynucleotide; an organism transformed with such an expression vector; and methods of synthesizing such an inhibitor and of inhibiting a serine proteinase.

BACKGROUND OF THE INVENTION

Serine Proteinases. Of the four major classes of proteinases (serine, cysteine, carboxyl, and metallo), serine proteinases, named for the serine residue at the active site, have been the most extensively studied. Serine proteinases fall into two families, subtilisin-like and chymotrypsin-like proteinases. As a general rule, the proteinases of these two families are not homologous: i.e., they have little similarity in amino acid sequence or structure with one another. However, within a family there is a considerable degree of similarity in three dimensional structure. Despite the differences between the serine proteinase families, all serine proteinases hydrolyze their peptide substrates according to the same mechanism.

The interaction between proteinases and their substrates is generally a "lock and key" interaction in which the substrate fits with steric precision into the active site of the proteinase. When a portion of the peptide substrate has fit into the proteinase active site, the enzyme's serine residue near the active site attacks a particular peptide bond of the substrate and

hydrolyzes it. After the hydrolysis, the product quickly departs from the proteinase.

In kinetic terms, a substrate with a good fit into the active site of a particular proteinase (or a particular enzyme having avidity for a substrate) has a low K_m. The rate of proteolysis, i.e., the instrinsic speed of proteolysis in a sample of enzyme and substrate is k_{cat}. The ratio of k_{cat}/K_m is considered to be a measure of an enzyme's catalytic efficiency.

The particular peptide bond of the peptide substrate which is hydrolyzed by a proteinase is generally referred to as the "scissile bond." A convention has been adopted (described in Schechter and Berger, *Biochem.Biophys.Res. Commun.* 27 157-162, 1967) to identify amino acid residues in a substrate with respect to the scissile bond: the two residues immediately N-terminal and C-terminal of the scissile bond are called P_1 and P_1' respectively. The remaining residues of the peptide substrate are P_2, P_3, P_4 etc. in sequential numbers increasing toward the N-terminal, and P_2', P_3', P_4' etc. in sequential numbers increasing toward the C-terminal. Using this convention, the four amino acid residues immediately upstream of the scissile bond are indicated P_4-P_3-P_2-P_1↓ (the ↓ indicating the scissile bond), or more simply P_4-P_3-P_2-P_1.

Studies of peptide substrates have revealed that proteinases generally require, in order for proteolysis to occur, one or more particular amino acids be present in close proximity to the scissile bond. The sequence of amino acids most commonly found in substrates of a particular proteinase is referred to as the "consensus sequence" for the substrates of that proteinase. Often, the consensus sequence residues come into contact with the active site of the enzyme and are located immediately upstream of the scissile bond.

Post-Translational Cleavage of Pro-Proteins. Numerous biologically active polypeptides, including peptide hormones, growth factors, certain blood clotting factors and envelope glycoproteins of viral pathogens, are initially synthesized as pro-proteins. These pro-proteins are often modified post-translationally by cleavage at the C-terminal side of two or more adjacent basic amino acids. For example, pro-insulin is cleaved at the C-terminal side of the consensus sequence Arg-Arg↓ or Lys-Arg↓ to form insulin. Numerous other pro-proteins that must be cleaved into an active form have been identified since the discovery of pro-insulin. However, the endoproteinase(s) effecting this cleavage *in vivo* in higher organisms were not discovered until recently.

One proteinase which has such proteolytic activity and has been known for years in eucaryotes is kexin, synthesized by the yeast *Saccharomyces cerevisiae*. Kexin is a subtilisin-like, calcium dependant, membrane bound endoproteinase having 814 amino acids (as opposed to the 382 amino acid residues of subtilisin BPN'). Kexin, which has multiple domains (prepro; subtilisin-like catalytic; "P"; middle; cysteine-rich; serine/threonine rich; trans-membrane; amphipathic alpha; and cytoplasmic domains), cleaves yeast proteins of the secretory pathway. It has a highly specific endoproteolytic activity, cleaving peptides at the sequences KR↓ and RR↓, but not at the sequence KK↓. Kexin also correctly processes pro-insulin expressed in yeast cells, pro-albumin, POMC and the precursor for protein C.

Recently, an endoproteinase having proteolytic activity similar to that of kexin was discovered in higher organisms. The endoproteinase, named furin (molecular weight around 90 kD), is found in most tissue types. Furin has been found to cleave several natural pro-proteins including pro-von Willebrand's factor, pro-β-nerve growth factor (van de Ven et al., *Crit.Revs. Oncogenesis*, 4 115-136 (1993); factor IX; anthrax pathogenic agent PA; and even a pro-furin, Molloy et al., *J.Biol.Chem.* 267, 16396-16402

(1992). These pro-proteins are cleaved at the sequences RSKR↓SL, RSKR↓SS, RKKR↓ and RTKR↓ respectively. A broad consensus has been reached that the sequence RXKR↓ is the optimal sequence for furin. However, further analysis indicates that the sequence RXXR↓ is sufficient for efficient processing of at least some furin substrates. Molloy et al., *supra*. The basic residue at P_4, in particular an R residue, appears to enhance the cleavage efficiency of furin, van de Ven et al., *Molecular Biol. Reports*, 14, 265-275 (1987). In typical SPC's, KR↓ and RR↓ appear to be the principal determinants of specificity.

Furin has been found in so many tissue types that it is believed to be ubiquitous in all mammalian tissues. While furin may not be the sole endoproteinase having activity comparable to that of kexin, it may be the primary enzyme responsible for post-translational processing in higher eucaryotes. Several other endoproteinases resembling furin have been discovered; these are however usually found in only one or a few tissues of mammals. Furin-like enzymes have also been found in lower organisms and insects.

In addition to playing a role in the maturation of many endogenous pro-proteins, furin is believed necessary to process the viral envelope glycoproteins encoded by viral polynucleotides to the size employed in the viral coat. Cleavage of glycoprotein gp 160 to a heterodimer composed of gp120 and gp41 is an essential step in HIV proliferation. According to Hallenberger et al., *Nature* 360, 358-361 (1992) human furin is the enzyme responsible for this cleavage.

Furin also may play a role in disease: although only limited data are available, elevated expression levels of furin are commonly seen in primary human non-small cell lung carcinomas, adenocarcinomas and squamous cell carcinomas.

Study of the amino acid sequences of kexin and furin have revealed consider-able similarity between them. Kexin and furin both contain essential alpha-helix and beta-sheet secondary structure elements that are characteristic of the protein fold of the subtilisin family. As noted, both furin and kexin have numerous domains. Three domains of furin and kexin have substantial amino acid sequence similarity. Furin is believed to have a trans-membrane and cytoplasmic domain, like kexin, which anchors the proteinase in the cytoplasmic membrane.

Due to the similarity in the primary and secondary structures of these two endoproteinases, one may employ molecular modelling to approximate the structure of furin from that of subtilisin. Computer-assisted molecular modelling to date predicts an increase in the number of negatively charged side chains on the substrate binding face of furin relative to subtilisin. It is believed that this high density of negative charge helps attract and perhaps orient highly positively charged substrate segments to the furin active site. van de Ven et al., *Crit.Revs.Oncogenesis*, *supra*, at 123.

The 90kD form of furin is believed to be a membrane bound form of the enzyme. However another form of furin has been found secreted in certain cell cultures. This "truncated furin," having molecular weight 75 kD, is believed to be unbound to any membrane, seems to retain the proteolytic activity of the full furin molecule and to lack the putative transmembrane and cytoplasmic domains.

Furin-like enzymes have been found in limited tissues and lower organisms include: PC2 found in human, rat and mouse cortex hypothalmus and spinal cord tissue; PACE4 in mouse osteosarcoma; PC4 in testes; *Xenopos* (PC2); *Drosophila* (Dfurin and Dfurin2); *Caenorhabiditis elegans* (encoded by the *bil-4* gene); and in hydra (a PC3-like enzyme). Together with furin and kexin, these enzymes are termed the subtilisin-like pro-protein convertases.

Protein Inhibitors of Serine Proteinases. The protein inhibitors of proteinases are commonly found in all organisms and are believed to play an important role in preventing undesirable proteolysis. The proteolysis which is thus inhibited may be from endogenous proteinases (i.e., synthesized by the same organism which synthesizes the inhibitor) or from exogenous proteinases. For example, secretory trypsin inhibitors in the pancreas prevent premature activation of endogenous trypsinogen and in turn other pancreatic zymogens; and avian ovomucoids, some investigators believe, inhibit the exogenous proteinases of bacteria which most commonly infect avian eggs.

Structurally, protein inhibitors of proteinases vary widely, as is reflected in the fact that ten families of inhibitor have been recognized. The primary factor defining each family is extensive amino acid sequence homology. Two further important factors are the conserved placement of disulfide bridges and the conserved location of the "reactive site peptide bond."

The interaction of an inhibitor with the active site of an enzyme is one which resembles the interaction between a proteinase and its substrate. The steric fit of the inhibitor's "reactive site peptide bond" and its surrounding area into the active site region involves numerous forces, including not only van der Waals and hydrogen bonding interactions, but also salt bridges.

The majority of known serine proteinase inhibitors act according to the "standard mechanism." These inhibitors, generally having 25 or more amino acids in poly-peptide sequence, contain on their surface a "reactive site peptide bond." This bond corresponds in several respects to the scissile bond of a protein substrate: it is a peptide bond which is attacked and is hydrolyzed (reversibly) by the proteinase. The scissile bond is located on a prominent exterior position of the inhibitor. The reactive site peptide bond

is found on the surface of peptide inhibitors, and is usually present in an oligopeptide loop within the protein. This loop is often formed by a disulfide bond or strong non-covalant bonds; consequently, hydrolysis of the reactive site peptide bond does not cleave two inhibitor halves from one another; the two halves remain together. This non-separation of the "modified" inhibitor's fragments reflects an important facet of the standard mechanism model: i.e., that the "modified" inhibitors also function as inhibitors by forming a complex with the proteinase and allowing re-formation of the original peptide bond. This ability of the modified inhibitor to inhibit reflects the equilibrium which exists between the enzyme and virgin inhibitors, the complex, and enzyme and modified inhibitors.

When the reactive site peptide bond and the surrounding region of a standard mechanism type inhibitor contact the active site of the proteinase, an enzyme-substrate complex is formed. The complex may disassociate either into the enzyme and the original inhibitor (referred to in the art as a "virgin" inhibitor) or into the enzyme and the inhibitor in which the reactive site peptide bond has been hydrolyzed (a "modified" inhibitor). However, a further important facet of the standard mechanism type inhibitor is that although the reactive site peptide bond is cleaved, this cleavage occurs only in the complex, and is slow. In kinetic terms, the values for k_{cat} and K_m for the cleavage of standard mechanism inhibitor are very unfavorable; however, the value of the k_{cat}/K_m ratio is in the range of values for normal substrates. Thus, hydrolysis of the reactive site peptide bond is extremely slow.

In light of these similarities to the scissile bond, the amino acid residues of peptide inhibitors are also identified with respect to the reactive site peptide bond using the abbreviations P_3, P_2, P_1, ↓, P_1', P_2', and P_3'.

Protein inhibitors have high specificity to proteinases. Replacing a single amino acid residue in the inhibitor near the reactive site peptide bond

may severely reduce the inhibitory activity. The P_1 residue generally plays the most important role in determining which proteinases are inhibited as well as the degree of inhibition. Yet protein inhibitors may absorb significant amino acid replacement even at key residue sites. In most bioactive proteins such as enzymes, the residues at structurally and functionally important positions are nearly unvaried while numerous changes are tolerated at structurally and functionally neutral positions. By contrast, protein inhibitors, which have similarly precise specificity to their target proteinases, are nevertheless able to tolerate replacement of an amino acid at the P_1 residue. Such a replacement may simply shift the original inhibitory activity from one proteinase to a different proteinase.

At present, it is unclear why proteinase inhibitors act as inhibitors instead of as substrates. In fact, a peptide inhibitor which inhibits the proteinase of one species may be a substrate for the corresponding enzyme in another species. Thus, strong inhibitors of bovine trypsin are good substrates for a homologous trypsin *1* from the starfish *Dermasterish imbricata*. Laskowski and Kato, *Ann.Rev.Biochem.* 49, 593-626 (1980).

Ovomucoids and Turkey Ovomucoid Third Domain Peptides. The ovomucoid proteins have been among the most extensively studied protein inhibitors of protein-ases. One reason for this is the ready availability of ovomucoids in relatively high amounts: they are one of the more abundant proteins of avian egg white, present at 10% of total protein, about 10 g/L. Another reason is that, since the turn of the century, it has been known that chicken ovomucoid is the major component respon-sible for trypsin inhibition by egg white.

Avian ovomucoids consist of three tandem, homologous domains which are closely homologous to pancreatic secretory trypsin inhibitor (of the Kazal family). Thus, the ovomucoids belong to the Kazal family of

inhibitors. Third domain ovo-mucoid peptides have been isolated and sequenced from 153 species of birds.

Despite the ready availability of ovomucoids in high volumes, the use of the full ovomucoid molecule can be somewhat cumbersome. This is due not only to the size of the molecule but to the fact that the first and second domains frequently inhibit a proteinase different from the third domain. These three domains maybe be cleaved without intra-domain nicking, the third domain being easiest to obtain (either by limited proteolysis or CNBr cleavage or a combination of both). Thus, the third domain of avian ovomucoid peptides has become one of the most studied protein inhibitors.

In fact, the third domain of turkey ovomucoid ("OMTKY3") is the ovomucoid third domain which is most studied. Sequencing studies of the avian ovomucoid third domains have shown that each residue in the turkey third domain has the amino acid most commonly found at that point in all other 152 species' third domains (and that the amino acid sequence of OMTKY3 is identical to that in the ovomucoid third domain of the ocellated turkey *Agriocharis ocellata*.) Laskowski et al. *Biochemistry* 26, 202-221 (1987); Laskowski et al., *Journal of Protein Chemistry*, 9. 715-725 (1990); 12, 419-433 (1993).

The different avian ovomucoid third domain peptides are perhaps the leading example of amino acid replacement at a key amino acid residue resulting not in inactivation but instead in bioactivity retention or change of specificity. This is seen in the fact that the third domain of silver pheasant ovomucoid, having Met at P_1, inhibits chymotrypsin, elastase and subtilisin, while the third domain of Japanese quail ovomucoid, having Lys at P_1, inhibits trypsin but none of the above proteinases.

OMTKY3 has 51 amino acid residues. Of these 51 amino acids, 40 are strongly conserved, being present in 130 of the 153 species' third

domains (85%). The remaining 11 residues are all strongly variable. Of these 11, 7 (i.e., residues 15, 17, 18, 20, 21, 32 and 36) contact the proteinases human leukocyte, elastase, bovine chymotrypsin, and SGPB, and therefore constitute a consensus contact residue set. Clearly, the contact residues are highly variable. (Residues 13, 14, 16, 19, and 33 also contact these proteinases.) The P_1 residue, amino acid 18, is so highly variable among the avian species that it is considered hypervariable.

Although hypervariable, only 11 out of the possible 20 L-amino acids have been found at the P_1 residue in the 153 species' third domain peptides. P_1 is never an aromatic amino acid in these peptides, but is most commonly a hydrophobic aliphatic (Leu and Met); P_1 is a charged amino acid (Lys) only once.

Modification of Protein Inhibitors of Serine Proteinases. It would be useful to have one or more protein inhibitors of serine proteinases such as furin for several reasons. Such inhibitors permit one to study the proteolytic mechanism of such proteinases and, *in vivo* or in cell cultures, to study the physiological role of such enzymes, particularly in diseases where elevated levels of furin are present. Since these inhibitors block the processing of viral coat proteins, they provide useful tools for studying viral replication processes, assembly and maturation processes. They are also useful as chemotherapeutic agents for *in vivo* treatment of viral infections and conditions associated with viral infections, including but not limited to HIV infections.

Several investigators have synthesized protein inhibitors of proteinases by modifying natural ones. Wieczorek et al., *Bioch.Biophys.Res.Comm.* 144 499-504 1987 described semi-synthetic hybrids of OMTKY3. The first hybrid linked amino acids 19 through 56 from natural OMTKY3 molecules to a synthetic oligopeptide having the sequence of amino acids 1 through 18 of OMTKY3; the second linked the same

natural fragment of OMTKY3 to a synthetic oligopeptide having the sequence of amino acids 6-18 of OMTKY3. The two semi-synthetic OMTKY3 hybrids were said to have the same equilibrium constant as natural OMTKY3.

Wolfson et al., *Biochemistry* 32, 5327-5331, 1993, reported replacing a tight turn in the amino acid sequence of cytokine human interleukin 1 beta with a loop from α_1-antitrypsin, an inhibitor of elastase. The resulting chimeric AT/IL peptide was said to retain interleukin activity but also to inhibit elastase.

International Application WO 90/10649 ("*Sambrook et al.*") describes the introduction of replacement amino acids into protein inhibitors of proteinases by DNA mutation for the purpose of increasing their inhibitory efficiency. The only modified inhibitors actually taught are serpin family inhibitors already known to inhibit t-pa. The mutations which *Sambrook et al.* disclose reverse the effect of a mutation introduced into the cognate endoproteinase (e.g., a residue in the serpin inhibitor PAI-1 is modified from E to R in order to counteract the mutation of an E residue in the t-pa active site).

Wild type OMTKY3 is a powerful inhibitor of many serine proteinases which exhibit a preference for neutral residues at P_1. Yet the specificity of OMTKY3 may be altered by replacing its P_1 residue. Native OMTKY3 is ineffective as an inhibitor of glutamic acid-specific *S. griseus* proteinase (GluSGP); however, Komiyama et al. *J. Biol.Chem.* 266 10727-10730 (1991) is said to be OMTKY3 converted into a powerful inhibitor by an L18E replacement (Glu replacing wild type Leu) at P_1. Similarly, although native OMTKY3 is ineffective against trypsin, an L18K mutation (Lys replacing wild type Leu) at P_1 is said to convert it into a moderately good trypsin inhibitor.

To date no protein inhibitors of furin, or of furin-like enzymes, which appear to be naturally directed against the proteinase have been identified. OMTKY3 itself does not inhibit furin. Two small molecular inhibitors, decanoyl-Arg-Ala-Lys-Arg-chloromethylketone and decanoyl-Arg-Gln-Lys-Arg-chloromethylketone are said to inhibit furin's endoproteinase activity. Stieneke-Grober et al., *EMBO J.* 11, 2407-2412 (1992).

SUMMARY OF THE INVENTION

Applicants have realized that inhibitors of furin, and of other subtilisin-like serine proteinases, may be made by introducing into known inhibitors the consensus sequence from substrates of furin, or other subtilisin-like serine proteinases.

In a first embodiment, the invention comprises a modified standard mechanism type protein proteinase inhibitor of a serine proteinase which hydrolyzes a target peptide at a scissile bond adjacent to a consensus sequence. The modified inhibitor comprises at least 25 amino acids, a reactive site peptide bond and a mutated amino acid sequence corresponding to the consensus sequence of the target peptide. This consensus sequence is located immediately adjacent to the reactive site peptide bond. The modified inhibitor may be a protein proteinase inhibitor of the Kazal family and the serine proteinase may be selected from the group consisting of kexin, furin, truncated furin, PC2, PC1/PC3, PC4, PACE-4, Dfurin1 and Dfurin2.

The modified inhibitor may comprise a mutated amino acid sequence selected from the group consisting of B_4-X_3-X_2-B_1, B_4-X_3-B_2-B_1 or B_2-B_1, where B is a basic amino acid selected from the group consisting of Arg or Lys; and X is any amino acid. The modified inhibitor incorporating this mutation may suitably be selected from the group consisting of mammalian pancreatic secretory trypsin inhibitor, including but not limited to human pancreatic secretory trypsin inhibitor; avian ovomucoid and avian ovomucoid

third domain. Suitable avian ovomucoid third domains include the turkey ovomucoid third domain in its virgin or modified state.

One suitable modified avian ovomucoid third domain peptide contains the mutated amino acid sequence B_4-X_3-B_2-B_1, wherein B_4 is R, B_2 is K and B_1 is R. The amino acid sequence of this particular inhibitor may further comprise the mutation selected from the group consisting of B_6, Y_1' and Z_3', wherein B is as defined above, Y is L or S and Z is R. Another suitable modified avian ovomucoid third domain peptide wherein the mutated amino acid sequence is selected from the group consisting of R_4-X_3-X_2-R_1 and B_2-R_1.

Still more particularly, the modified avian ovomucoid third domain peptide inhibitor may be selected from the group consisting of (A15R-T17K-L18R) OMTKY3; (K13R-A15R-T17K-L18R) OMTKY3; (A15R-T17K-L18R-E19S) OMTKY3; (A15R-T17K-L18R-R21L) OMTKY3; (T17K-L18R) OMTKY3; (T17R-L18R) OMTKY3; and (A15R-L18R) OMTKY3. There is further provided a composition comprising from 10 to 10,000 μg of one of these modified OMTKY3 inhibitors.

All of these inhibitors described above incorporating a consensus sequence from a target peptide have an inhibitory effect on those serine proteinases which cleave that target peptide. Thus, all these inhibitors are useful, as they may be used as laboratory reagents to study the serine proteinases or as chemotherapeutic agents to treat diseases associated with these proteinases.

In another embodiment of the invention, there is provided a purified, isolated polynucleotide comprising one or more sequences of nucleotide bases which collectively encode one of the modified standard mechanism type inhibitors including a mutated amino acid sequence selected from the group consisting B_4-X_3-X_2-B_1, B_4-X_3-B_2-B_1 or B_2-B_1, where B is a basic amino

acid selected from the group consisting of Arg or Lys; and X is any amino acid.

The polynucleotide may suitably encode the amino acid sequence for a modified inhibitor selected from the group consisting of a mammalian pancreatic secretory trypsin inhibitor, including but not limited to human pancreatic secretory trypsin inhibitor; avian ovomucoid and avian ovomucoid third domain. The polynucleotide may also comprise one or more sequences of nucleotide bases which collectively encode a modified avian ovomucoid third domain peptide comprising one of the mutated amino acid sequences described above, e.g., a modified peptide selected from the group consisting of (A15R-T17K-L18R) OMTKY3; (K13R-A15R-T17K-L18R) OMTKY3; (A15R-T17K-L18R-E19S) OMTKY3; (A15R-T17K-L18R-R21L) OMTKY3; (T17K-L18R) OMTKY3; (T17R-L18R) OMTKY3; and (A15R-L18R) OMTKY3.

The polynucleotide may alternatively comprise one or more sequences of nucleotide bases which are antiparallel and complementary to one of the above polynucleotides. Other suitable polynucleotides may comprise one or more sequences of nucleotide bases which are substantially homologous to the nucleotide base sequence in one of the polynucleotides encoding an OMTKY3 peptide inhibitor. Substantial homology between two polynucleotide strands is present where there is base pairing between 75% or more of the two strands' bases.

Still more particularly this embodiment comprises a polydeoxyribonucleotide comprising a nucleotide sequence of 153 nucleotide bases. These nucleotide bases correspond to basis 696 through 848 of pEZZ318.tky2. However, in this polydeoxyribonucleotide, the sequence corresponding to nucleotide bases 717 through 737 are selected from the group consisting of

GTGAGTACCCTcgtCCTcgtTGtAaaCgtGAATACAGACCTCTCTGTG;
GTGAGTACCCTAAGCCTcgtTGtAaaCgttctTACAGACCTCTCTGTG;

GTGAGTACCCTAAGCCTcgtTGtAaaCgtGAATACctACCTCTCTGTG;
GTGAGTACCCTAAGCCTGCATGtAaaCgtGAATACAGACCTCTCTGTG;
GTGAGTACCCTAAGCCTGCATGtcgtCgtGAATACAGACCTCTCTGTG;

and

GTGAGTACCCTAAGCCTcgtTGtACGCgtGAATACAGACCTCTCTGTG.

These polynucleotides have numerous uses. They may be employed to generate an mRNA with point mutations, and thence a modified inhibitor peptide with a mutated amino acid sequence. Alternatively, they may be used as probes specific for polynucleotides or polynucleotide sequences.

In further embodiment of the invention, there is provided an expression vector comprising one of the above polydeoxyribonucleotides. Suitably, such an expression vector may further comprise one or more sequences of nucleotide bases which collectively confer resistance to an antibiotic upon an organism. This embodiment of the invention further provides for a polydeoxyribonucleotide wherein the one or more sequences of bases collectively encode the modified amino acid sequence of a modified turkey ovomucoid third domain peptide and further comprises, 5' or 3' of said sequence of bases collectively encoding said modified peptide, one or more appropriate regulatory control sequences which collectively enable expression of polydeoxyribonucleotide.

The expression vector may further comprise one or more sequences of nuc-leotide bases which collectively encode a second peptide such that the expression product of said vector is a fusion protein. One suitable expression vector comprises the nucleotide sequence of pEZZ318.tky2. All the above expression vectors may be made according to techniques known in the art.

These expression vectors may be employed to synthesize the modified inhibitor peptides, either by *in vitro* synthesis or by transforming a micro-organism with the vector and culturing the vector under conditions

favorable to growth of said organism and to expression of the vector in order to generate large amounts of the inhibitor.

In a further embodiment of the invention, there is provided an organism transformed with one of the above expression vectors. The organism which is transformed may be any organism which is conventionally transformed with such vectors. One suitable organism is one which has been transformed with an expression vector comprising the nucleotide sequence of PEZZ318.tky2. In particular, one suitable transformed organism is *Escherichia coli* RV308 (ATCC accession number 31608). The organism transformed with an expression vector may be produced by methods known in the art. Such transformed organisms may be used to synthesize large amounts of the modified inhibitor.

A further embodiment of the invention is a fusion protein expressed from one of the above expression vectors. A fusion protein includes the amino acid sequence of the modified standard mechanism type inhibitor of a serine proteinase fused to a "secondary protein." The secondary protein in a fusion protein may impart several advantages, e.g., to simplify purification of the modified inhibitor from cell lysates. Fusion proteins containing staphylococcal Protein A may easily be removed from a cell extract by affinity chromatography. The secondary protein may also help target and direct the modified inhibitor to certain tissue types. Thus, the secondary protein may be an antibody specific to the target tissue or a molecule which interacts specifically with a receptor present on the surface of the target cells. Where the target tissue is virally infected, the secondary protein may suitably be an antibody specific to a non-neutralizing epitope on a virus; an envelope protein which is cleaved by a subtilisin-like enzyme such as furin. The secondary protein may also be a moiety which is taken up by the target cells, as by endocytosis, thus bringing the entire fusion protein into the target cell. Such secondary proteins may be selected by

persons skilled in the art. Fusion proteins may be prepared according to techniques known in the art.

The fusion proteins may be used either as to supply the modified inhibitor, ob-tained by cleaving away the superfluous peptide; or to deliver the modified inhibitor specifically to tissues where one desires to inhibit proteinase activity.

A further embodiment of the invention is a method of synthesizing a modified turkey ovomucoid domain peptide, comprising the steps of transforming an organism with an expression vector comprising a nucleotide sequence including pEZZ318.tky2; culturing the transformed organism for one or more generations under conditions favorable to growth of the transformed organism and to expression of the expression vector; and isolating the modified turkey ovomucoid third domain peptide by lysing the progeny of said cultured transformed organism to form a cell-free extract, and isolating the peptide from the extract. This embodiment has a self-evident utility.

There is also provided a method of inhibiting proteolytic activity of a subtilisin-like serine proteinase comprising the steps of synthesizing a Kazal group inhibitor having a mutated amino acid sequence selected from the group consisting of B_4-X_3-X_2-B_1, B_4-X_3-B_2-B_1 or B_2-B_1 and exposing the subtilisin-like serine proteinase to said modified inhibitor. B is a basic amino acid selected from the group consisting of Arg or Lys; and X is any amino acid. In this method, the subtilisin-like serine proteinase may be selected from the group consisting of kexin, furin, truncated furin, PC2, PC1/PC3, PC4, PACE-4, Dfurin1 and Dfurin2. The Kazal inhibitor may be selected from the group consisting of (A15R-T17K-L18R) OMTKY3; (K13R-A15R-T17K-L18R) OMTKY3; (A15R-T17K-L18R-E19S) OMTKY3; (A15R-T17K-L18R-R21L) OMTKY3; (T17K-L18R) OMTKY3; (T17R-L18R) OMTKY3; and

(A15R-L18R) OMTKY3. The steps in all the above methods may be practiced according to techniques known in the art.

This method is useful for exploring the mechanism of proteolysis, the physiological role of such proteinases, or for inhibiting such proteinases *in vivo* to treat an illness, such as one caused by a virus requiring such proteinase activity.

Finally, there is provided a kit for measuring the inhibition of proteinase activity comprising at least one of the modified turkey ovomucoid third domain peptide inhibitors. Such a kit may be designed according to rules well known in the art. This embodiment has a self-evident utility.

It is understood that in some technical areas, such as chemical inventions relating to catalysis, the mechanism underlying an invention is unknown and the resulting unpredictability of the invention may require extensive experimental data to demonstrate the utility of the invention as claimed. The instant invention, which relates to inhibitors of enzymatic proteinases, might be considered analogous to catalytic inventions since it concerns modifying the amino acid sequence of peptides in order to modify their interaction with biological catalysts. However, this invention is not based on the mechanism of catalysis, but rather on the highly precise steric fit between proteinases and their substrates or inhibitors. Applicants have discovered that incorporating certain mutated amino acid sequences (the consensus sequence from a "target peptide," defined below) into protein inhibitors at certain locations (adjacent to the reactive site peptide bond) imparts inhibitory activity to protein inhibitors which before had no such activity against a certain serine proteinase. This modification of the protein inhibitors allows part of the inhibitor to fit sterically into the proteinase active site. Consequently, the present invention is much less unpredictable than catalysis inventions, and extensive experimental support for the invention is not required.

It is further recognized that claims having broad scope also might obligate one to submit extensive experimental support to demonstrate that the full genus claimed is in fact operative. Certain claims of the instant application are phrased in broad terms, i.e., making a specific modification in a broad range of protein inhibitors. The broad terms actually describe a narrow and specific modification of protein inhibitors. The inhibitors to be modified are specifically identified as those having no inhibitory activity against a certain serine proteinase. The modification is specific: mutating an inhibitor's amino acid sequence to that of a consensus sequence. The position of the modification in the inhibitor is specific: it is made adjacent to the reactive site peptide bond. The breadth of the claims herein is due to the wide applicability of the modification: persons skilled in the art would now expect this modification to incorporate into a wide range of protein inhibitors, along with the consensus sequence, an ability to interact specifically with proteinases which cleave target peptides containing that sequence. In light of the specificity of the claims and the view of those skilled in the art, extensive experimental support of the claims is not required.

Upon review of the art discussed above, it might at first appear that the modifications of known protein inhibitors of the instant invention are taught in the art. Replacement of P_1 in OMTKY3 has been seen to impart inhibitory activity against proteinases which native OMTKY3 does not inhibit; and decanoyl-Arg-Ala-Lys-Arg-chloro-methylketone, which includes the R-X-K-R consensus sequence of certain furin sub-strates, is said to inhibit the endoproteinase activity of furin. Some of the native avian ovomucoid third domain amino acid sequences among the 153 species tested have Lys at P_1, Arg at P_2, as well as Arg at P_6 and Leu at P_3'.

However, the replacement of the P_1 residue alone has been known as a way to alter the specificity of protein inhibitors. Thus, the replacement of P_1 in OMTKY3 does not indicate that one could replace one or more

further amino acids near the reactive site peptide bond in OMTKY3 and thereby impart inhibitory activity against subtilisin-like pro-protein convertases. The inhibition from the chloromethylketone results from a rather small molecule. One skilled in the art could not expect that this molecule's tetrapeptide sequence (R-A-K-R) incorporated into a protein inhibitor, having at least 25 amino acids, would retain the same inhibitory activity. This is because one could not know whether the tetrapeptide's configuration in the inhibitor would be the same as in tetrapeptide. Finally, the mere presence of certain amino acid residues at P_1, P_2, P_6 and P_3' in separate native avian ovomucoids of species other than turkey provides no motivation to replace the amino acids at these residues in OMTKY3, much less to make the additional replacement of P_4. Thus, upon closer inspection, the art does not lead one to modify OMTKY3 by incorporating as a mutated amino acid sequence the consensus sequence of serine proteinase target peptides.

On the contrary, there are strong indications which discourage so modifying OMTKY3. It is difficult to find inhibitors in which P_4 and P_1 are Arg. Only one inhibitor among the Kazal domain sequences has been found having P_4 as Arg. The chicken ovoinhibitor fourth domain has the sequence RRCPR,I where the "," indicating the reactive site peptide bond. Scott et al., *J.Biol.Chem.* 262 5899-5907 (1987). In the squash inhibitor family however, Arg at P_4, and the sequence R-X-X-R preceding the reactive site peptide bond are not uncommon. The water melon (*Citrullus vulgaris*) trypsin inhibitor I has the same sequence as chicken ovoinhibitor fourth domain. Otlewski, et al., *Biol.Chem. Hoppe-Seyler* 368 1505-1507 (1987). Neither of these protein inhibitors is effective as an inhibitor of furin. Thus, even if one skilled in the art considered incorporating the consensus sequence of subtilisin-like pro-protein convertase substrates into OMTKY3 in order to inhibit these convertases, the inactivity of naturally available peptides already incorporating this sequence against these convertases would discourage one from carrying out this modification.

DESCRIPTION OF THE DRAWINGS

Fig.1 illustrates the amino acid sequence of the 6-56 OMTKY3 (wild type), its reactive site peptide bond (↓), likely chain folding and disulfide bonds placement.

Fig. 2 illustrates the amino acid sequence and the likely chain folding and disulfide bonds of the wild type OMTKY3 of Fig. 1 with one mutation: L18R, i.e., original Leu^{18} is replaced by Arg^{18}.

Fig. 3 illustrates the amino acid sequence, and the likely chain folding and disulfide bonds of the wild type OMTKY3 of Fig. 1 with three mutations: A15R, T17K and L18R, i.e., original Ala ^{15} is replaced with Arg^{15}; original Thr ^{17} with Lys^{17}; and original Leu^{18} with Arg^{18}.

Fig. 4 illustrates the procedure for constructing plasmid pEZZ318.tky2, which encodes (A15R-T17K-L18R) OMTKY3.

Fig. 5 is a restriction map of plasmid pEZZ318.tky2.

Fig. 6 is a graph illustrating the effect of the (A15R-T17K-L18R) OMTKY3 on free furin concentration. The curve drawn through the points is the best fit for K_a- $1.1 \times 10^7 M^{-1}$. pH is 7.5 and temp is 30°C. The molar concentrations of the total inhibitor present were measured directly but the molar concentrations of free furin were obtained from these data by an iterative process. The dashed line is what would be expected if the inhibition were stoichiometric.

DETAILED DESCRIPTION OF THE INVENTION

Nomenclature.

The nomenclature used to define the peptides is that specified by the IUPAC-IUB Commissioner on Biochemical Nomenclature (*European J. Biochem.*, 1984, 138, 9-37). By natural amino acid is meant one of the

common, naturally occurring amino acids found in proteins: Gly, Ala, Val, Leu, Ile, Ser, Thr, Lys, Arg, Asp, Asn, Glu, Gln, Cys, Met, Phe, Tyr, Pro, Trp and His. When the amino acid residue has isomeric forms, it is the L-form of the amino acid that is represented unless otherwise expressly indicated. These three letter abbreviations for amino acids are employed herein, except in describing a point mutation, in which case the following single letter amino acid abbreviations, corresponding respectively to the three letter abbreviations above, are employed: G, A, V, L, I, S, T, K, R, D, N, E, Q, C, M, F, Y, P, W and H.

Other abbreviations used are:

Boc	tert-butyloxycarbonyl-
BOP	benzotriazole-1-yl-oxy-tris-(dimethylamino)phosphonium hexafluorophosphate
Cbz	benzyloxycarbonyl
2-Cl-Cbz	2-chloro-benzyloxycarbonyl-
DCB	2,6 dichlorobenzyl-
DCCI	dicyclohexylcarbodiimide
DIC	diisopropylcarbodiimide
DCM	dichloromethane
DMF	dimethylformamide
Fmoc-	fluorenylmethyloxycarbonyl
HOBt	1-hydroxybenzotriazole
HPLC	high peformance liquid chromatography
MeOH	methyl alcohol
Tos	p-toluensulfonyl-
TEA	triethylamine
TFA	trifluoroacetic acid
tBu	tert-butyl

The convention under which peptides are written with the amino terminal placed to the left and the carboxy terminal to the right is followed

herein. A further convention of numbering peptide residues, starting with residue one at the N-terminal and proceeding sequentially to the C-terminal residue, is followed herein, with an exception for avian ovomucoid third domain peptides. In all the avian ovomucoid third domain peptides (except those in the Sequence Listing), the amino acid residues are numbered from N-terminal residue six (Val^6) through C-terminal 56 (Cys^{56}). This convention arose from the use in early studies of a ovomucoid fragments including the five residue long peptide joining the second and third domains. This connecting peptide of avian ovomucoid third domain peptides (residues 1 through 5) has since been found to have no effect on the analog's interaction with enzymes (Wieczorek et al., *Biochem.Biophys.Res.Comm.* 144 494-504 (1987). Accordingly, investigators now studying the inhibitory effect of the ovomucoid third domain cleave the connecting peptide (residues 1-5) away and study only what is considered to be the third domain, residues 6-56.

A further convention defining amino acid replacements induced by mutation is followed herein. Under this convention, the amino acid mutation is indicated by a single letter preceding and following a number: "L18R." The number indicates a particular amino acid residue in the subject peptide which has been altered. The letters are single letter amino acid abbreviations, the first being the amino acid at the numbered residue in non-mutated peptide, the second being the replacement amino acid. Thus, (A15R-T17K-L18R) OMTKY3 represents three mutations in a turkey ovomucoid third domain peptide: Ala^{15} is replaced by Arg, Thr^{17} by Lys, and Leu^{18} by Arg. (Although residues "15," "17" and "18" of this peptide are actually the tenth, twelfth and thirteenth residues from the N-terminal, they are numbered according to the convention discussed in the preceding paragraph because the peptide is a avian ovomucoid third domain.)

All polynucleotide sequences discussed herein may be either single or double stranded, unless indicated otherwise; further, they may be

comprised of RNA or DNA, or DNA-RNA double stranded hybrid. The DNA sequences set forth herein follow standard conventions; that is, sequences are written from 5' on the left to 3' on the right (except where two base pairing sequences are written parallel to one another: then the lower strand is written 3' to 5' to illustrate the inter-strand base pairs). Certain oligonucleotide primer sequences used to induce substitution mutations are divided into triplet codons to facilitate identifying which codon is mutated.

Under a further convention followed herein, the term "upstream" as applied to peptides and polynucleotides is understood to mean a direction or position which is toward the N-terminal or 5' direction respectively. Conversely, "downstream" is understood to mean, in peptides and polynucleotides, a direction or position which is toward the C-terminal or 3' direction respectively.

A convenient abbreviation employed herein combines the P_4-P_3-P_2-P_1 terms (identifying amino acid residues by their position respective to the scissile or reactive site peptide bond) with the terms "B" and "X," defined above. This abbreviation employs "B" or "X" in the former term to specify the type of amino acid residue which is present at a certain residue. Thus, "B_2-B_1" is understood to mean that P_2 and P_1 are B, i.e., a basic amino acid selected from the group consisting of Arg and Lys. Similarly, B_4-X_3-B_2-B_1 is understood to mean that the residue at P_4, P_2 and P_1 are all B, while the P_3 residue may be any amino acid.

The term "target peptide" is understood herein to mean proteinaceous compounds which have one of their peptide bonds cleaved by the proteinase during a specific interaction with a proteinase active site. Two suitable target peptides include peptides which are natural substrates of a proteinase as well as other peptides which, although not encountering the proteinase in nature, still are cleaved by it.

The Preferred Embodiments.

The first embodiment of the invention provides modified "standard mechanism" type inhibitors. These may be synthesized by suitable methods, including but not limited to exclusively or partially solid phase techniques, fragment condensation, classical solution phase synthesis or recombinant DNA techniques. Suitable solid phase synthesis is generally described by Merrifield, *J.Am.Chem.Soc.*, 85, p. 2149 (1963); suitable recombinant DNA techniques are described in Maniatis et al., Molecular Cloning, A Laboratory Manual (2nd Ed.) Cold Spring Harbor Press, 1989. Other equivalent chemical synthesis methods known in the art may also be used.

A suitable method of synthesizing modified "standard mechanism" type inhibitors, such as the Kazal family inhibitors, in general, or modified turkey ovo-mucoid third domain peptides in particular is to express the inhibitor from a plasmid genetically encoding it. This may be accomplished using the polynucleotides making the second embodiment of the invention.

A cDNA segment encoding the native inhibitor protein to be modified is isolated, either by excising the desired cDNA segment from a readily available plasmid by restriction endonuclease cleavage; or, if a cDNA segment encoding the native inhibitor is not available, by isolating a cDNA segment which encodes a similar inhibitor protein. In the case of turkey ovomucoid third domain, one selects a cDNA encoding chicken ovomucoid third domain, as for example by excising the desired cDNA segment using appropriate endonuclease from a plasmid carrying such a segment.

The isolated cDNA segment is then incorporated into a double stranded plasmid. One selects the double stranded plasmid from those which also have cleav-age sites for the same restriction endonuclease at suitable positions, and treats the double stranded plasmid with the appropriate endonuclease(s) to generate the partially single stranded base

sequences complementary to those of the isolated cDNA segment. The isolated cDNA segment is then added to the cleaved plasmid, and the two are allowed to anneal. In this manner, the cDNA encoding chicken ovomucoid third domain is incorporated into plasmid pEZZ318 (in Fig. 4). The plasmid bearing the inhibitor sequence (e.g. p318OM3D.chi1) is ligated and replicated in single stranded form by helper virus.

Deletion and Substitution Mutations. The necessary mutations are then introduced into the plasmid so that it encodes a serine proteinase inhibitor with a mutated amino acid sequence immediately adjacent to its reactive site peptide bond which corresponds to a consensus sequence one of the proteinase's substrates. If one did not initially isolate a cDNA segment for the desired native inhibitor, but instead isolated a cDNA segment encoding a similar inhibitor protein, further mutations are also introduced into the plasmid to convert cDNA sequence into one encoding the desired native inhibitor (described below). These mutations need not be performed in any particular order.

Where one wishes to generate the modified standard mechanism inhibitor as part of a fusion protein, one selects a plasmid which also encodes for a suitable secondary protein. Certain proteins, referred to as "marker" proteins, enable one to purify the fusion protein from a cell extract simply. One such marker protein is staphylococcal Protein A; fusion proteins containing Protein A may easily be removed from a cell extract by affinity chromatography. The cDNA segment encoding the inhibitor is inserted adjacent to and in reading frame with the segment coding for this protein moiety to assure their co-transcription and co-translation as a fusion protein. Furthermore, if one desires to isolate the modified inhibitor after expression from the remainder of the fusion protein, there should be means for separating the two peptide segments. This may be done by introducing one or more codons between the nucleotide bases encoding one or more amino acid which make the resulting expressed fusion protein vulnerable to

chemical or protease cleavage. Suitably, a Met residue renders a fusion protein cleavable by CNBr treatment. If necessary, extraneous segments should also be removed from the plasmid by well-known deletion mutation techniques. When these steps are performed on plasmid pEZZ318, the resulting plasmid is p318OM3D.chi3, which codes for a Protein A-inhibitor fusion protein.

If necessary, this plasmid may then be then subjected to substitution mutation to convert the DNA sequence from that of the isolated inhibitor peptide into that of the desired native type. Substitution mutations in the chicken ovomucoid third domain convert the sequence into one encoding for OMTKY3 in plasmid pEZZ318.tky1.

Finally, the necessary codons are altered by substitution mutation to introduce into the native standard mechanism type inhibitor, immediately adjacent to its reactive site peptide bond, a consensus sequence from a substrate of the target serine proteinase. For example, substitution mutations may be introduced into certain codons of plasmid pEZZ318.tky1 to introduce the desired BXBB, BXXB or BB amino acid sequences so that the segment encodes for (A15R, T17K and L18R) OMTKY3. The results of these last mutations appear in plasmid pEZZ318.tky2 (sequence below and restriction map in Fig. 5).

The deletion mutations are introduced by well known methods. Thus for example, a partially homologous single strand oligonucleotide is annealed to the single stranded plasmid encoding both the marker and the ovomucoid third domain protein. This oligonucleotide's base sequence is homologous to base sequences adjacent to the 5' and 3' sides of the segment to be deleted; however, it has no sequence homology to the intervening segment. Annealing the oligonucleotide to the single stranded plasmid results in base pairing between the oligonucleotide and corresponding homologous sequences on the plasmid. As a result of this pairing, the two base paired

sequences on the plasmid are drawn toward one another, with the intervening single-stranded segment forming a single stranded loop. A DNA polymerase is added to replicate the plasmid. Polymerization initiates at the primer and produces two circles of unequal length: an original, full length strand and a "copied" strand including the oligonucleotide with of course the undesired segment absent. The two unequal circles are then segregated by conventional techniques, as via replication and molecular cloning.

The substitution mutations are also performed by well known techniques. Thus, for example, an oligonucleotide is annealed to the single stranded plasmid DNA. Unlike the oligonucleotide used in deletion mutations, this oligonucleotide is fully homologous to an uninterrupted target sequence on the plasmid, except for mis-pairings at from about one to several bases. Despite this mis-pairing, the oligonucleotide and the plasmid hybridize under conditions of moderate stringency since almost all of the oligonucleotide's bases pair with homologous plasmid bases. With this partial double stranding, polymerization may be initiated. The single stranded plasmid is replicated by a DNA polymerase. The original single stranded plasmid of course remains unchanged; however, the new copy bears the mutated codons of the primer. These replacements in one or more codons, thus effect substitution mutations.

Upon replication and cloning into individual colonies, the two different plasmids are identified by conventional screening procedures using for example probing with labeled mutant specific oligonucleotides and/or DNA base sequencing.

Expression of the Plasmid. Once the final plasmid -- e.g. a plasmid bearing the marker protein- (A15R-T17K-L18R) OMTKY3 fusion protein -- has been generated, it is used to transform an appropriate expression host. Any antibiotic resistance genes which are present on the plasmid exert selection pressure toward retention of the plasmid. The transformed

expression host provides a permanently reproducible source of the plasmid, but even more importantly of the desired fusion protein. The fusion protein is isolated and subjected to chemical or proteinase cleavage to yield the desired OMTKY3 analog.

Demonstration of Furin Inhibition. The inhibitory effect of the modified standard mechanism type inhibitor on the target serine proteinase may be tested as follows. In a first assay, a fixed amount of the serine proteinase is mixed in a cuvette with a peptide compound which fluoresces when cleaved. After an appropriate incubation period, fluorescence is measured. Then a sample of the modified standard mechanism type inhibitor is added to the cuvette and, after a further appropriate incubation time, fluorescence is again measured. This assay is repeated several times, each time using the same concentration of serine proteinase and fluorescing compound, however with a varying inhibitor concentration. For each assay, the level of uninhibited and inhibited fluorescence are plotted against time.

From each of these plots, one calculates the concentration of free serine proteinase, i.e., the concentration of proteinase unbound by the inhibitor, from the slope of the curve on the plot. One then plots the concentration of free serine proteinase against the varying concentrations of inhibitor. Such a plot is seen Fig. 4; it demonstrates that as the concentration of inhibitor rises, the concentration of free serine proteinase falls. This type of relationship clearly indicates inhibitory activity on the part of the inhibitor.

EXAMPLE I

(A15R-T17K-L18R) OMTKY3 is built step by step on a benzhydrylamine HCl resin containing about 1 mEq NH_2/g (Advanced ChemTech, Louisville, KY) in a reaction vessel for manual solid-phase synthesis starting with Boc-Cys in accordance with the procedures set forth below.

The benzhydrylamine HCl resin (1 g, about 1 mmol), after neutralization with 10% TEA in CH_2Cl_2, is coupled sequentially with a 3 molar excess of protected amino acid in accordance with the Schedule as follows:

STEP	REAGENTS AND OPERATIONS	MIXING TIMES (min)
1	Coupling: Boc-amino acid in DCM or DMF depending on the solubility of the particular protected amino acid, plus DIC	60-90
2	MeOH (or DMF then MeOH) wash	2
3	DCM wash	2
4	MeOH wash	2
5	DCM wash (three times)	2
6	Deprotection: 50% TFA in DCM (twice)	5 and 25
7	DCM wash	2
8	2-Propanol wash	1
9	Neutralization: 10% TEA in DCM	2
10	MeOH wash	1
11	Neutralization: 10% TEA in DCM	2
12	MeOH wash	1
13	DCM wash (three times)	2

After attachment of Boc-Cys to the resin, the following amino acids are added sequentially to the resin according to the above schedule: Boc-Lys[Z(2-Cl)][55], Boc-Gly, Boc-Phe, Boc-His(Tos), Boc-Ser(Bzl), Boc-Leu[50], Boc-Thr, Boc-Leu, Boc-Thr, Boc-Gly, Boc-Asn[45], Boc-Ser(Bzl), Boc-Glu, Boc-Val, Boc-Val, Boc-Ala[40], Boc-Asn, Boc-Cys, Boc-Phe, Boc-Asn, Boc-Cys[35], Boc-Lys[Z(2-Cl)], Boc-Asn, Boc-Gly, Boc-Tyr, Boc-Thr[30], Boc-Lys[Z(2-Cl)], Boc-Asn, Boc-Asp, Boc-Ser(Bzl), Boc-Gly[25], Boc-Cys, Boc-Leu, Boc-Pro, Boc-

Arg(Tos), Boc-Tyr[20], Boc-Glu, Boc-Arg(Tos), Boc-Lys[Z(2-Cl)], Boc-Cys, Boc-Arg(Tos),[15] Boc-Pro, Boc-Lys[Z(2-Cl)], Boc-Pro, Boc-Tyr, Boc-Glu,[10] Boc-Ser(Bzl), Boc-Cys, Boc-Asp, and Boc-Val[6].

The resulting resin-oligopeptide, which includes all the amino acids of (A15R-T17K-L18R) OMTKY3, is treated with 50-fold excess acetic anhydride and TEA in 30 ml DMF for 30 min. The acetylated peptide-resin is then washed with DMF (3 times), iPrOH (3 times) and DCM (3 times) and dried in vacuo. Removal of the protecting groups and cleavage of the decapeptide from the resin is carried out by

with 2 ml anisole and 20 ml of HF at 0° for 45 min. After elimination of HF under vacuum, the peptide-resin remainder is washed with dry diethyl ether. The peptide is then extracted with 50% aqueous acetic acid, separated from the resin by filtration, and lyophilized.

Crude peptides (860 mg, 725 mg) are purified on Column A with solvent system i using a linear gradient of 10-40 % B in 60 min at flow rate of 30 ml/min. 230 nm. Purified peptides prove to be substantially (>96%) pure in analytical HPLC by using solvent system i in a linear gradient mode (15-35%B in 20 min). Amino acid analysis gives the expected results.

EXAMPLE II

Preparation of Plasmid pEZZ318.tky2 encoding OMTKY3 Analog

1. Construction of p318OM3D.chi1. The steps for construction of the vector for the expression of chicken ovomucoid third domain as a downstream fusion segment to IgG binding domains of staphylococcal Protein A are shown in Figure 4. The plasmid pOM100 encoding the chicken ovomucoid cDNA described in O'Malley et al., *Cell* 18 (1979) 829-42, is treated with *Hin* C II/*Bam*H I to release a double-stranded segment encoding the chicken ovomucoid third domain. This fragment is purified by gel electrophoresis and inserted into a 4.6kbp plasmid, pEZZ318, (described

in Nilsson et al., *1992 Current Opin. in Structural Biol.* 2 569-75) already digested by *Sma* I/*Bam*H I in-frame to the coding sequences for the IgG binding ("Z") domains of staphylococcal protein A. The resulting double-stranded plasmid is replicated in *E. coli* using M13K07 helper phage in accord with Viera and Messing, *Methods in Enzymol.* 154 382-403, 1987 to obtain a single-stranded template.

2. Conversion of p318OM3D.chi1 to p318OM3D.chi2. The conversion to p318OM3D.chi2 is accomplished by deletion mutagenesis in accord with Nakamaye and Eckstein, *Nucleic Acids* 14 9679-9698 (1986) using the following "protein A/OM3D fusion primer":

5' GCA GTC AAC CAT AGC GAA GTG AGC TTC TTT AGC 3'

This primer hybridizes to part of the "Z" and "OM3D" single strand sequences but not to the intervening segment. As noted above, all DNA primers herein are written from 5' left to 3' right by convention. Thus, reading the plasmid 5' to 3' starting at base 714, it is clear the primer is substantially homologous thereto.

Polymerization is initiated from this double stranded portion of the plasmid with Klenow fragment and dNTP's, then the newly synthesized strand ends are ligated with T4 DNA ligase. The resultant mutant plasmid p318OM3D.chi2 is isolated as described in Nakame and Eckstein, *supra* and *Maniatis et al., supra.*

3. Conversion of p318OM3D.chi2 to p318OM3D.chi3. The p318OM3D.chi2 plasmid is subjected to deletion mutagenesis by digesting it with *Bam*H I and *Hin*D III in standard reaction conditions. The resulting mixture of large and small segments is separated by gel electrophoresis. The large fragment of plasmid p318OM3D.chi2 is then incubated with Klenow fragment and dNTP's to render its single stranded terminal portions double stranded; these double stranded ends are then ligated with T4 DNA ligase. The resulting plasmid is p318OM3D.chi3.

4. Mutation of p318OM3D.chi3 to pEZZ318.tky1. The chicken ovomucoid third domain sequence encoded in p318OM3D.chi3 is mutated to the turkey ovomucoid by site-directed mutagenesis. The double-stranded p318OM3D.chi3 plasmid is replicated in *E. coli* using M13K07 helper phage in accord with Viera and Messing, *supra*, to obtain a single-stranded template. A single-stranded template DNA using the "CHI>TKY" oligonucleotide primer is added:

5' GTT GTC GGA TCC ACA GAG AGG TCT GTA TTC CAG CGT GCA TGC AGG CTT AGG 3'

This primer replaces the original nucleotides with those that are underlined. The primer not only alters the encoded ovomucoid third domain from chicken to turkey; it also introduces a unique *Bam*H I site by silent mutation in the codon for Gly25. This mutation is also underlined. The codon changes GAC>GCA; GCA>CTG; and GAC>TAC produce the following amino acid changes D15A; A18L; and D20Y respectively. The resultant expression plasmid is termed pEZZ318.tky1.

5. Conversion of pEZZ318.tky1 to pEZZ318.tky2. This conversion is carried out by a further substitution mutation, using an oligonucleotide designed to mutate the codons for P4, P2, and P1. This oligonucleotide the following sequence:

5' GTT GTC GGA GCC ACA GAG AGG TCT GTA TTC TCG TTT GCA TCG AGG CTT AGG -3'

Mutated bases are underscored. At position 10 from the 5' end of the oligonucleotide, there is an additional mutation which eliminates a unique *Bam*H I site of p318EZZ.tky1. This mutation is silent, but is used to facilitate screening for mutants.

6. Expression and purification of wild type turkey ovomucoid third domain and its analogs. Organisms transformed with plasmid pEZZ318.tky2 are inoculated into a medium containing ampicillin. Only transformed organisms which retain the ampicillin resistance encoded on the

plasmid pEZZ318.tky2 may grow. This culture is grown for 12 or more hours and the periplasmic protein A / (A15R-T17K-L18R) OMTKY3 fusion protein is isolated by conventional steps. Suitably, one applies osmotic shock to the expression host to release the periplasmic proteins (Randall and Hardy 1986 *Cell* 46 921-928), centrifuges the shocked cells and subjects the supernatant to affinity chromatography to obtain the fusion protein. The fusion protein is subjected to CNBr cleavage, and the resulting (A15R-T17K-L18R) OMTKY3 is isolated and purified by size exclusion and ion exchange chromatography.

EXAMPLE III

Expression of (A15R-T17K-L18R) OMTKY3 from *E. coli* Strain RV308

A variation of the procedure of Nilsson and Abrahmson, *supra*, may be used to express the staphylococcal protein A - (A15R-T17K-L18R) OMTKY3 fusion protein: an overnight culture of *E. coli* strain RV308 transformed as in Example II is grown at 37°C in LB broth (10 g/liter Tryptone [Difco], 5 g/liter Yeast Extract [Difco], 10 g/liter NaCl) to which ampicillin (final concentration 250 mg/liter), and glucose (0.1%) have been added. This is diluted 100-1000 times and grown 18 hours at 37°C in a baffled flask. An osmotic shock procedure (according to Randall and Hardy, *supra*, is used to obtain periplasmic proteins. Cells from a one liter culture are obtained by cooling to 0°C and centrifuging at 8.000 x g at 4°C for 20 min. The cells are suspended in 50 ml of ice cold sucrose buffer (0.5 M sucrose, 0.1 M Tris-HCl, pH 8.2, 1 mM EDTA) and placed in an ice-water bath for 10 min. To this is added 0.80 ml of an egg white lysozyme solution (10 mg/ml in water), and immediately 50 ml of ice-cold water is added. The mixture is kept on ice for 5 min. and then 1.8 ml of I M $MgSO_4$ added. The osmotically shocked cells are centrifuged at 8,000 x g at 4°C for 20 min.

The supernatant was considered to be the periplasmic fraction and is filtered through glass fiber filter paper (GF/D - Whatman). Affinity

chromatography per Nilsson and Abrahmson, *supra* is carried out using IgG Sepharose 6 Fast Flow gel (Pharmacia). The column is washed with ten bed volumes of TST (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, and 0.05% Tween 20), the periplasmic fraction loaded, and the column washed with an additional 20 bed volumes of TST, then with ten bed volumes of 1 mM NH_4Ac. The protein is eluted with three bed volumes of 0.5 M acetic acid titrated to pH 3.3 using NH_4Ac.

The Protein A domain is split off by CNBr cleavage of the linker peptide (the only Met in the fusion protein) and isolated by size exclusion and ion exchange chromatography. The (A15R-T17K-L18R) OMTKY3 analog is characterized by amino acid analysis, sequencing (Proton 2020) from residue 6 through 19 and by mass spectrometry (Westec) with electrospray injection. All of the analyses, sequences and molecular masses are in accord with expectations.

EXAMPLE IV

Measuring Inhibition of Furin Proteolysis by (A15R-T17K-L18R) OMTKY3

Furin used in measuring Inhibition. As noted above, furin has a carboxy terminal transmembrane domain. In this assay, a COOH terminally truncated human fragment developed and used for detailed studies on furin in vitro (Molloy et al., *supra* is employed to assay proteinase inhibition. This molecule, which terminates at residue 713 (Leu) and lacks the transmembrane domain, is believed to have the same protease activity as the complete furin.

This truncated furin is expressed in African Green monkey kidney cells BSC-40 from a vaccinia virus recombinant VV:hFUR713t as described by Molloy et al. *supra*. Final purification of the truncated furin is on Pharmacia MonoQ 5/5 anion exchange column: the eluted 250 μl fractions are monitored by a furin enzymatic assay using a fluorogenic peptide

substrate and SDS-PAGE. Three fractions constituting the peak of activity serve as the furin source.

The Fluorescence Assays. The molar concentration of the OMTKY3 analog in the suspension is measured by injecting aliquots of the suspension onto a TSK G2000 analytical size exclusion column. The effluent is monitored at 206nm by an LKB detector and the area under the inhibitor peak integrated. This is compared to the area under an equal size aliquot of standardized OMTKY3 of equal volume.

The activity of human furin is determined by monitoring the hydrolysis of 1.2×10^{-4}M (final concentration) of t-butyloxycarbonylarginyl-valyl-arginylarginyl-4-methylcoumarin-7-amide. The enzyme and inhibitor are incubated together in the reaction medium for 2 hr to reach equilibrium. Then substrate is added and substrate hydrolysis allowed to proceed for 16 hrs. The hydrolysis is stopped by addition of a 60-fold excess of 1mM $ZnCl_2$ solution. The reactions are conducted in 0.1M HEPES, 0.5% Triton-X-100, 0.00IM $CaCl_2$, 0.00IM β-mercaptoethanol, pH 7.50 at 30°C. The fluorescence of these solutions is then measured at 460nm (10nm slit width) on a Perkin Elmer LS50 spectrofluorometer using 370nm wavelength (10nm slit width) for excitation.

The stock solutions of substrate are made in DMSO to avoid hydrolysis during storage. Only small volumes of these solutions are added (generally 10μl to 3 mls of reaction mixture) to avoid perturbation effects.

A blank of buffer + substrate is also measured. The concentration of furin is then calculated from the formula

$$[E_x] = E_0 \times S_x / S_0$$

where $[E_x]$ is the free enzyme concentration in the reaction mixture, E_0 is the unvarying concentration of total furin, S_x is the slope of the absorbance vs time plot (in absorbance units per second) and S_0 is the slope of the

uninhibited run. The results are then submitted to a nonlinear least square curve fitting program described in the Ph.D. thesis of Richard Wynn submitted to the faculty of Purdue University, May, 1991 (pp. 28-29). The resulting concentration of free furin is then plotted against the concentrations of (A15R-T17K-L18R) OMTKY3 in the different assays.

Addition of massive amounts of wild type OMTKY3 or of (L18R) OMTKY3 produces insignificant inhibition of furin activity. We estimate that K_a's for these inhibitors are $200M^{-1}$ and $450M^{-1}$ respectively. However, the K_a from (A15R-T17K-L18R) OMTKY3 is found to be $1.1 \times 10^7 M^{-1}$ and clear inhibition was observed (Fig. 6).

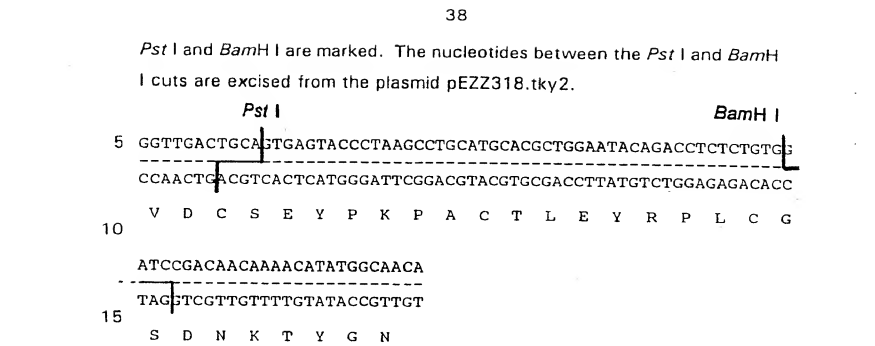
Fitting of these data caused some problems. Normally the operational molarity of the enzymes is measured prior to the experiment either by burst titration or by titration with a very strong inhibitor. Neither is available for human furin. Therefore, an iterative fitting procedure was employed to yield the molar furin concentrations given on the vertical axis of Figure 6; s can be seen, the fit in Fig. 6 is quite good.

EXAMPLE V

Cassette mutagenesis of turkey ovomucoid third domain is performed on double-stranded plasmid DNA of the phagemid pEZZ318TKY-MET, containing OM3TKY fused to the C-terminus of two protein A domains. The plasmid DNA is digested overnight at 37°CD with the restriction endonucleases *Bam*H I (GibcoBRL) and *Pst* I (GibcoBRL) in React3 Buffer (GibcoBRL). The digested DNA is run on an agarose gel (1% agarose in TAE) and the appropriate band is excised from the gel and purified by Geneclean ™ method of DNA-gel slice purification (Bio101).

The wild type coding sequence for turkey ovomucoid third domain as ex-pressed in pEZZ318.tky2 appears below (with the protein translation appearing in one letter code below the sequence). The sites of cleavage for

Pst I and *Bam*H I are marked. The nucleotides between the *Pst* I and *Bam*H I cuts are excised from the plasmid pEZZ318.tky2.



Six gapped duplex DNA cassettes to introduce the mutations are made from synthetic oligonucleotides. The oligonucleotides, synthesized on an Applied Biosciences DNA synthesizer, have the following sequences (in 5' to 3' order) to cause the indicated mutations in the peptide sequence of OM3TKY:

K13R-A15R-T17K-L18R:

GTGAGTACCCTcgtCCTcgtTGtAaaCgtGAATACAGACCTCTCTGTG;

A15R-T17K-L18R-E19S:

GTGAGTACCCTAAGCCTcgtTGtAaaCgttctTACAGACCTCTCTGTG;

A15R-T17K-L18R-R21L:

GTGAGTACCCTAAGCCTcgtTGtAaaCgtGAATACctACCTCTCTGTG;

T17K-L18R:

GTGAGTACCCTAAGCCTGCATGtAaaCgtGAATACAGACCTCTCTGTG;

T17R-L18R:

GTGAGTACCCTAAGCCTGCATGtcgtCgtGAATACAGACCTCTCTGTG;

and

A15R-L18R:

GTGAGTACCCTAAGCCTcgtTGtACGCgtGAATACAGACCTCTCTGTG.

At position 23 in these oligonucleotides, there is a silent mutation (C to T) which eliminates a unique *Sph* I site. This mutation is used for restriction enzyme selection of mutants.

Two oligonucleotides, designated *Pst* I-Overlap and *Bam*H I-Overlap, are designed to overlap the ends of the mutagenic oligonucleotide and also either *Pst* I or *Bam*H I, respectively. These two oligonucleotides (both in 5' to 3' orientation) are:

Pst I - Overlap	AGGGTACTCAACTGCA
*Bam*H I-Overlap	GATCCACAGAGAGG

After synthesis, each oligonucleotide is manually deprotected in ammonium hydroxide, and run on a 12% polyacrylamide/urea gel for purification. The appro-priate sized bands are excised from the gel, and the DNA eluted overnight in "crush and soak" buffer (*Maniatis et al., supra*). The resulting eluate is run over a Whatman GC C18 column for final purification.

After purification, all the oligonucleotides are phosphorylated with T4 polynucleotide kinase (New England Biolabs, Beverly, MA). Equimolar amounts of phosphorylated *Pst* I Overlap, *Bam*H I-Overlap, and a single mutagenic oligonucleotide are mixed together, and heated to 65°C. These mixtures are then allowed gradually to cool to room temperature, and are used as a cassette in a ligation reaction. The ligation reactions are done in ligase buffer (*Maniatis et al., supra*) with T4 DNA ligase (NEB) at 16°C overnight.

The following six mutagenesis cassettes are formed from ligation of the above oligonucleotides. The top oligonuclotide is written 5′ to 3′ and contains the mutations marked in lower case letters. The *Pst* I overlap and *Bam*H I overlap primers appear on the second line of each set in appropriate 3′ to 5′ orientation. These two oligonucleotides should anneal to the ends of the mutagenic oligonucleotide, and also to the sticky ends of the *Pst* I and *Bam*H I sites.

K13R-A15R-T17K-L18R:

GTGAGTACCCTcgtCCTcgtTGtAaaCgtGAATACAGACCTCTCTGTG
ACGTCACTCATGGGA GGAGAGACACCTAG

A15R-T17K-L18R-E19S:

GTGAGTACCCTAAGCCTcgtTGtAaaCgttctTACAGACCTCTCTGTG
ACGTCACTCATGGGA GGAGAGACACCTAG

A15R-T17K-L18R-R21L:

GTGAGTACCCTAAGCCTcgtTGtAaaCgtGAATACctACCTCTCTGTG
ACGTCACTCATGGGA GGAGAGACACCTAG

T17K-L18R:

GTGAGTACCCTAAGCCTGCATGtAaaCgtGAATACAGACCTCTCTGTG
ACGTCACTCATGGGA GGAGAGACACCTAG

T17R-L18R:

GTGAGTACCCTAAGCCTGCATGtcgtCgtGAATACAGACCTCTCTGTG
ACGTCACTCATGGGA GGAGAGACACCTAG

A15R-L18R:

GTGAGTACCCTAAGCCTcgtTGtACGCgtGAATACAGACCTCTCTGTG
ACGTCACTCATGGGA GGAGAGACACCTAG

The ligation reactions are transformed into XL1-Blue cells and grown with ampicillin and tetracycline selection. Double stranded DNA is prepared by the alkaline-lysis plasmid preparation method and subsequently digested with *Sph* I and transformed into XL1-Blue cells. Colonies from these transformations are screened for *Sph* I resistance. The *Sph* I resistant samples are sequenced by the dideoxy method and the correct mutant DNAs are transformed into *E. coli* RV308 cells for protein expression.

The DNA sequence of plasmid pEZZ318.tky2 follows. Also shown are the amino acids for the described fusion protein (including a pre-sequence for Protein A). The amino acids encoded by the plasmid appear from base 192 to base 848, with base 192 to base 299 encoding the secretory pre-peptide of Protein A, base 300 to base 671 encoding the IgG binding domains of staphylococcal Protein A, base 672 to base 695 encoding the linker peptide region, and base 696 to base 848 encoding the OMTKY3 analog. The arrow marks the site of CNBr cleavage of the fusion protein.

Although the invention has been described with regard to its preferred embodiments, it should be understood that changes and modifications known to one having the ordinary skill in this art may be made without departing from the scope of the invention, which is set forth in the claims which are appended thereto. Substitutions known in the art which do not significantly detract from its effectiveness may be employed in the invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Anderson, Stephen
Laskowski Jr., Michael V.

(ii) TITLE OF INVENTION: NOVEL PROTEIN INHIBITORS OF SERINE PROTEINASES

(iii) NUMBER OF SEQUENCES: 20

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Omri M. Behr, Esq.
(B) STREET: 325 Pierson Avenue
(C) CITY: Edison
(D) STATE: NJ
(E) COUNTRY: USA
(F) ZIP: 08837

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/089,248
(B) FILING DATE: 07-JUL-1993

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Behr, Omri M.
(B) REGISTRATION NUMBER: 22,940
(C) REFERENCE/DOCKET NUMBER: RUTG3.0-019

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (908) 494-5240
(B) TELEFAX: (908) 494-0428
(C) TELEX: BEPATEDIN 5

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4800 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCGGCCGCTC GAAATAGCGT GATTTTGCGG TTTTAAGCCT TTTACTTCCT GAATAAATCT 60

TTCAGCAAAA TATTTATTTT ATAAGTTGTA AAACTTACCT TTAAATTTAA TTATAAATAT 120

AGATTTTAGT ATTGCAATAC ATAATTCGTT ATATTATGAT GACTTTACAA ATACATACAG 180

GGGGTATTAA TTTGAAAAAG AAAAACATTT ATTCAATTCG TAAACTAGGT GTAGGTATTG 240
CATCTGTAAC TTTAGGTACA TTACTTATAT CTGGTGGCGT AACACCTGCT GCAAATGCTG 300
CGCAACACGA TGAAGCCGTA GACAACAAAT TCAACAAAGA ACAACAAAAC GCGTTCTATG 360
AGATCTTACA TTTACCTAAC TTAAACGAAG AACAACGAAA CGCCTTCATC CAAAGTTTAA 420
AAGATGACCC AAGCCAAAGC GCTAACCTTT TAGCAGAAGC TAAAAAGCTA AATGATGCTC 480
AGGCGCCGAA AGTAGACAAC AAATTCAACA AAGAACAACA AAACGCGTTC TATGAGATCT 540
TACATTTACC TAACTTAAAC GAAGAACAAC GAAACGCCTT CATCCAAAGT TTAAAAGATG 600
ACCCAAGCCA AAGCGCTAAC CTTTTAGCAG AAGCTAAAAA GCTAAATGAT GCTCAGGCGC 660
CGAAAGTAGA CCGTAAAGAA GCTCACTTCG CTATGGTTGA CTGCAGTGAG TACCCTAAGC 720
CTCGATGCAA ACGAGAATAC AGACCTCTCT GTGGCTCCGA CAACAAAACA TATGGCAACA 780
AGTGCAACTT CTGCAATGCA GTCGTGGAAA GCAACGGGAC TCTCACTTTA AGCCATTTTG 840
GAAAATGCTG AATATCAGAG CTGAGAGAAT TCACCACAGG ATCAGCTTGG CACTGGCCGT 900
CGTTTTACAA CGTACTGACT GGGAAAACCC TGGCGTTACC CAACTTAATC GCCTTGCAGC 960
ACATCCCCCC TTCGCCAGCT GGCGTAATAG CGAAGAGGCC CGCACCGATC GCCCTTCCCA 1020
ACAGTTGCGT AGCCTGAATG GCGAATAATT CCAGACGATT GAGCGTCAAA ATGTAGGTAT 1080
TTCCATGAGC GTTTTTCCTG TTGCAATGGC TGGCGGTAAT ATTGTTCTGG ATATTACCAG 1140
CAAGGCCGAT AGTTTGAGTT CTTCTACTCA GGCAAGTGAT GTTATTACTA ATCAAAGAAG 1200
TATTGCGACA ACGGTTAATT TGCGTGATGG ACAGACTCTT TTACTCGGTG GCCTCACTGA 1260
TTATAAAAAC ACTTCTCAGG ATTCTGGCGT ACCGTTCCTG TCTAAAATCC CTTTAATCGG 1320
CCTCCTGTTT AGCTCCCGCT CTGATTCTAA CGAGGAAAGC ACGTTATACG TGCTCGTCAA 1380
AGCAACCATA GTACGCGCCC TGTAGCGGCG CATTAAGCGC GGCGGGTGTG GTGGTTACGC 1440
GCAGCGTGAC CGCTACACTT GCCAGCGCCC TAGCGCCCGC TCCTTTCGCT TTCTTCCCTT 1500
CCTTTCTCGC CACGTTCGCC GGCTTTCCCC GTCAAGCTCT AAATCGGGGG CTCCCTTTAG 1560
GGTTCCGATT TAGTGCTTTA CGGCACCTCG ACCCCAAAAA ACTTGATTAG GGTGATGGTT 1620
CACGTAGTGG GCCATCGCCC TGATAGACGG TTTTTCGCCC TTTGACGTTG GAGTCCACGT 1680
TCTTTAATAG TGGACTCTTG TTCCAAACTG GAACAACACT CAACCCTATC TCGGTCTATT 1740
CTTTTGATTT ATAAGGGATT TTGCCGATTT CGGCCTATTG GTTAAAAAAT GAGCTGATTT 1800
AACAAAAATT TAACGCGAAT TTTAACAAAA TATTAACGTT TACAATTTAA ATATTTGCTT 1860
ATACAATCTT CCTGTTTTTG GGGCTTTTCT GATTATCAAC GGGGGTACAT ATGATTGACA 1920
TGCTAGTTTT ACGATTACCG TTCATCGATT CTCTTGTTTG CTCCAGACTC TCAGGCAATG 1980
ACCTGATAGC CTTTGTAGAC CTCTCAAAAA TAGCTACCCT CTCCGGCATG AATTTATCAG 2040
CTAGAACGGT TGAATATCAT ATTGATGGTG ATTTGACTGT CTCCGGCCTT TCTCACCCGT 2100
TTGAATCTTT ACCTACACAT TACTCAGGCA TTGCATTTAA AATATATGAG GGTTCTAAAA 2160
ATTTTTATCC TTGCGTTGAA ATAAAGGCTT CTCCCGCAAA AGTATTACAG GGTCATAATG 2220

TTTTTGGTAC AACCGATTTA GCTTTATGCT CTGAGGCTTT ATTGCTTAAT TTTGCTAATT 2280
CTTTGCCTTG CCTGTATGAT TTATTGGATG TTGGAATTTG ATGCGGTATT TTCTCCTTAC 2340
GCATCTGTGC GGTATTTCAC ACCGCATATG GTGCACTCTC AGTACAATCT GCTCTGATGC 2400
CGCATAGTTA AGCCAGCCCC GACACCCGCC AACACCCGCT GACGCGCCCT GACGGGCTTG 2460
TCTGCTCCCG GCATCCGCTT ACAGACAAGC TGTGACCGTC TCCGGGAGCT GCATGTGTCA 2520
GAGGTTTTCA CCGTCATCAC CGAAACGCGC GAGGCAGCTT GAAGACGAAA GGGCCTCGTG 2580
ATACGCCTAT TTTTATAGGT TAATGTCATG ATAATAATGG TTTCTTAGAC GTCAGGTGGC 2640
ACTTTTCGGG GAAATGTGCG CGGAACCCCT ATTTGTTTAT TTTTCTAAAT ACATTCAAAT 2700
ATGTATCCGC TCATGAGACA ATAACCCTGA TAAATGCTTC AATAATATTG AAAAAGGAAG 2760
AGTATGAGTA TTCAACATTT CCGTGTCGCC CTTATTCCCT TTTTTGCGGC ATTTTGCCTT 2820
CCTGTTTTTG CTCACCCAGA AACGCTGGTG AAAGTAAAAG ATGCTGAAGA TCAGTTGGGT 2880
GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCTTGA GAGTTTTCGC 2940
CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTTAAAGTTC TGCTATGTGG CGCGGTATTA 3000
TCCCGTGTTG ACGCCGGGCA AGAGCAACTC GGTCGCCGCA TACACTATTC TCAGAATGAC 3060
TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG ATGGCATGAC AGTAAGAGAA 3120
TTATGCAGTG CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT TCTGACAACG 3180
ATCGGAGGAC CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA TGTAACTCGC 3240
CTTGATCGTT GGGAACCGGA GCTGAATGAA GCCATACCAA ACGACGAGCG TGACACCACG 3300
ATGCCTGTAG CAATGGCAAC AACGTTGCGC AAACTATTAA CTGGCGAACT ACTTACTCTA 3360
GCTTCCCGGC AACAATTAAT AGACTGGATG GAGGCGGATA AAGTTGCAGG ACCACTTCTG 3420
CGCTCGGCCC TTCCGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG TGAGCGTGGG 3480
TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT CGTAGTTATC 3540
TACACGACGG GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT 3600
GCCTCACTGA TTAAGCATTG GTAACTGTCA GACCAAGTTT ACTCATATAT ACTTTAGATT 3660
GATTTAAAAC TTCATTTTTA ATTTAAAAGG ATCTAGGTGA AGATCCTTTT TGATAATCTC 3720
ATGACCAAAA TCCCTTAACG TGAGTTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAAAG 3780
ATCAAAGGAT CTTCTTGAGA TCCTTTTTTT CTGCGCGTAA TCTGCTGCTT GCAAACAAAA 3840
AAACCACCGC TACCAGCGGT GGTTTGTTTG CCGGATCAAG AGCTACCAAC TCTTTTTCCG 3900
AAGGTAACTG GCTTCAGCAG AGCGCAGATA CCAAATACTG TCCTTCTAGT GTAGCCGTAG 3960
TTAGGCCACC ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT GCTAATCCTG 4020
TTACCAGTGG CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGGTTGGA CTCAAGACGA 4080
TAGTTACCGG ATAAGGCGCA GCGGTCGGGC TGAACGGGGG GTTCGTGCAC ACAGCCCAGC 4140
TTGGAGCGAA CGACCTACAC CGAACTGAGA TACCTACAGC GTGAGCATTG AGAAAGCGCC 4200
ACGCTTCCCG AAGGGAGAAA GGCGGACAGG TATCCGGTAA GCGGCAGGGT CGGAACAGGA 4260

GAGCGCACGA GGGAGCTTCA AGGGGGAAAC GCCTGGTATC TTTATAGTCC TGTCGGGTTT 4320

CGCCACCTCT GACTTGAGCG TCGATTTTTG TGATGCTCGT CAGGGGGGCG GAGCCTATGG 4380

AAAAACGCCA GCAACGCGGC CTTTTTACGG TTCCTGGCCT TTTGCTGGCC TTTTGCTCAC 4440

ATGTTCTTTC CTGCGTTATC CCCTGATTCT GTGGATAACC GTATTACCGC CTTTGAGTGA 4500

GCTGATACCG CTCGCCGCAG CCGAACGACC GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG 4560

GAAGAGCGCC CAATACGCAA ACCGCCTCTC CCCGCGCGTT GGCCGATTCA TTAATCCAGC 4620

TGGCACGACA GGTTTCCCGA CTGGAAAGCG GGCAGTGAGC GCAACGCAAT TAATGTGAGT 4680

TACCTCACTC ATTAGGCACC CCAGGCTTTA CACTTTATGC TTCCGGCTCG TATGTTGTGT 4740

GGAATTGTGA GCGGATAACA ATTTCACACA GGAAACAGCT ATGACCATGA TTACGAATTA 4800

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTGAGTACCC TCGTCCTCGT TGTAAACGTG AATACAGACC TCTCTGTG 48

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTGAGTACCC TAAGCCTCGT TGTAAACGTT CTTACAGACC TCTCTGTG 48

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTGAGTACCC TAAGCCTCGT TGTAAACGTT CTTACAGACC TCTCTGTG 48

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTGAGTACCC TAAGCCTCGT TGTAAACGTG AATACACTCC TCTCTGTG 48

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTGAGTACCC TAAGCCTGCA TGTAAACGTG AATACAGACC TCTCTGTG 48

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 49 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTGAGTACCC TAAGCCTGCA TGTCGTCGTG AATACAGACC TCTCTGTGA 49

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GTGAGTACCC TAAGCCTCGT TGTACGCGTG AATACAGACC TCTCTGTG 48

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCAGTCAACC ATAGCGAAGT GAGCTTCTTT AGC 33

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GTTGTCGGAT CCACAGAGAG GTCTGTATTC CAGCGTGCAT GCAGGCTTAG G 51

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AGGGTACTCA ACTGCA 16

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GATCCACAGA GAGG 14

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Val Asp Cys Ser Glu Tyr Pro Lys Pro Ala Cys Thr Leu Glu Tyr Arg
1               5                  10                  15

Pro Leu Cys Gly Ser Asp Asn Lys Thr Tyr Gly Asn Lys Cys Asn Phe
            20                  25                  30

Cys Asn Ala Val Val Glu Ser Asn Gly Thr Leu Thr Leu Ser His Phe
        35                  40                  45
```

Gly Lys Cys
50

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Val Asp Cys Ser Glu Tyr Pro Lys Pro Ala Cys Thr Arg Glu Tyr Arg
1 5 10 15

Pro Leu Cys Gly Ser Asp Asn Lys Thr Tyr Gly Asn Lys Cys Asn Phe
20 25 30

Cys Asn Ala Val Val Glu Ser Asn Gly Thr Leu Thr Leu Ser His Phe
35 40 45

Gly Lys Cys
50

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Val Asp Cys Ser Glu Tyr Pro Lys Pro Arg Cys Lys Arg Glu Tyr Arg
1 5 10 15

Pro Leu Cys Gly Ser Asp Asn Lys Thr Tyr Gly Asn Lys Cys Asn Phe
20 25 30

Cys Asn Ala Val Val Glu Ser Asn Gly Thr Leu Thr Leu Ser His Phe
35 40 45

Gly Lys Cys
50

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Arg Ala Lys Arg
1

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Arg Gln Lys Arg
1

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 86 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GGTTGACTGC AGTGAGTACC CTAAGCCTGC ATGCACGCTG GAATACAGAC CTCTCTGTGG 60

CCAACTGACG TCACTCATGG GATTCG 86

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 86 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TGTTGCCATA TGTTTTGTTG CTGGATACAA CGGTATACAA AACAACAGCC TACCACAGAG 60

AGGTCTGTAT TCCAGCGTGC ATGACG 86

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GTTGTCGGAG CCACAGAGAG GTCTGTATTC TCGTTTGCAT CGAGGCTTAG G 51

We claim:

1. A modified standard mechanism type protein proteinase inhibitor of a serine proteinase which hydrolyzes a target peptide at a scissile bond adjacent to a consensus sequence in said target peptide, said modified inhibitor comprising at least 25 amino acids, a reactive site peptide bond, and immediately adjacent to said reactive site peptide bond, a mutated amino acid sequence corresponding to said consensus sequence of said target peptide.

2. A modified inhibitor of a serine proteinase according to Claim 1, wherein said serine proteinase is a subtilisin-like proteinase.

3. A modified inhibitor of a serine proteinase according to Claim 2, wherein said inhibitor is a member of the Kazal family and said serine proteinase is selected from the group consisting of kexin, furin, truncated furin, PC2, PC1/PC3, PC4, PACE-4, Dfurin1 and Dfurin2.

4. A modified inhibitor according to Claim 3, wherein said mutated amino acid sequence is selected from the group consisting of B_4-X_3-X_2-B_1, B_4-X_3-B_2-B_1 or B_2-B_1, where B is a basic amino acid selected from the group consisting of Arg or Lys; and X is any amino acid.

5. A modified inhibitor according to Claim 4 which is selected from the group consisting of: pancreatic secretory trypsin inhibitor, avian ovomucoid and avian ovomucoid third domain.

6. An avian ovomucoid third domain according to Claim 5 which is selected from the group consisting of OMTKY3.

7. A modified turkey ovomucoid third domain peptide inhibitor according to Claim 6 wherein said mutated amino acid sequence is B_4-X_3-B_2-B_1, wherein B_4 is R, B_2 is K and B_1 is R.

8. A modified turkey ovomucoid third domain peptide inhibitor according to Claim 7 wherein said mutated amino acid sequence further comprises a mutation selected from the group consisting of B_6, Y_1', and Z_3', wherein B and R are as defined above, Y is L or S and Z is R.

9. A modified turkey ovomucoid third domain peptide inhibitor according to Claim 4 wherein said mutated amino acid sequence is selected from the group consisting of R_4-X_3-X_2-R_1 and B_2-R_1.

10. A modified turkey ovomucoid third domain peptide inhibitor according to Claim 4 comprising a mutated amino acid sequence, said inhibitor being selected from the group consisting of (A15R-T17K-L18R) OMTKY3; (K13R-A15R-T17K-L18R) OMTKY3; (A15R-T17K-L18R-E19S) OMTKY3; (A15R-T17K-L18R-R21L) OMTKY3; (T17K-L18R) OMTKY3; (T17R-L18R) OMTKY3; and (A15R-L18R) OMTKY3.

11. A composition comprising from 10 to 10,000 μg of one of said inhibitors according to Claim 10.

12. A polynucleotide comprising one or more sequences of nucleotide bases which collectively encode the amino acid sequence of one of said modified turkey ovomucoid third domain peptide inhibitors of Claim 10.

13. A polynucleotide comprising one or more sequences of nucleotide bases, which collectively are antiparallel and complementary to a polynucleotide of Claim 12.

14. A polynucleotide comprising one or more sequences of nucleotide bases which is substantially homologous to the nucleotide base sequence of a polynucleotide of Claim 12.

15. A polydeoxyribonucleotide according to Claim 12 comprising a nucleotide sequence of 153 nucleotide bases corresponding to bases 696 to 848 of pEZZ318.tky2, and in which the sequence corresponding to nucleotide bases 717 to 737 of said pEZZ318.tky2 is selected from the group consisting of

GTGAGTACCCTAAGCCTcgtTGtAaaCgtGAATACAGACCTCTCTGTG;
GTGAGTACCCTcgtCCTcgtTGtAaaCgtGAATACAGACCTCTCTGTG;
GTGAGTACCCTAAGCCTcgtTGtAaaCgttctTACAGACCTCTCTGTG;
GTGAGTACCCTAAGCCTcgtTGtAaaCgtGAATACctACCTCTCTGTG;
GTGAGTACCCTAAGCCTGCATGtAaaCgtGAATACAGACCTCTCTGTG;
GTGAGTACCCTAAGCCTGCATGtcgtCgtGAATACAGACCTCTCTGTG; and
GTGAGTACCCTAAGCCTcgtTGtACGCgtGAATACAGACCTCTCTGTG.

16. An expression vector comprising a polydeoxyribonucleotide of Claim 15.

17. An expression vector according to Claim 16 which further comprises one or more squences of nucleotide bases which collectively encode a second peptide such that the expression product of said vector is a fusion protein.

18. An expression vector according to Claim 17 comprising the nucleotide sequence of pEZZ318.tky2.

19. A transgenic organism transformed with an expression vector of Claim 16.

20. A transgenic organism transformed with an expression vector of Claim 18.

21. The transformed organism of Claim 20 which is *E. coli* RV308 (ATCC Accession No. 31608).

22. A fusion protein expressed from an expression vector of Claim 18.

23. A method of synthesizing a modified turkey ovomucoid third domain peptide of Claim 10, said method comprising

transforming an organism with an expression vector comprising a nucleotide sequence of pEZZ318.tky2;

culturing said transformed organism for one or more generations under conditions favorable to growth of said transformed organism and to expression of said expression vector; and

isolating turkey ovomucoid third domain peptide by lysing the progeny of said cultured transformed organism to form a cell-free extract, and isolating said peptide from said extract.

24. A method of inhibiting proteolytic activity of a subtilisin-like serine proteinase, said method comprising the steps of

synthesizing a Kazal group inhibitor having a mutated amino acid sequence selected from the group consisting of B_4-X_3-X_2-B_1, B_4-X_3-B_2-B_1 or B_2-B_1, where B is a basic amino acid selected from the group consisting of Arg or Lys; and X is any amino acid; and

exposing said subtilisin-like serine proteinase to said modified inhibitor.

25. A method according to Claim 24 wherein said subtilisin-like serine proteinase is selected from the group consisting of kexin, furin, truncated furin, PC2, PC1/PC3, PC4, PACE-4, Dfurin1 and Dfurin2.

26. A method according to Claim 25 wherein said Kazal inhibitor is selected from the group consisting of (A15R-T17K-L18R) OMTKY3; (K13R-

A15R-T17K-L18R) OMTKY3; (A15R-T17K-L18R-E19S) OMTKY3; (A15R-T17K-L18R-R21L) OMTKY3; (T17K-L18R) OMTKY3; (T17R-L18R) OMTKY3; and (A15R-L18R) OMTKY3.

27. A kit for measuring the inhibition of proteinase activity comprising at least one of the modified turkey ovomucoid third domain peptide inhibitors according to Claim 10.